Role of the Polycomb-Group Gene *FIE* in *Arabidopsis thaliana* Seed Development

Christine Stock

Doctor of Philosophy
The University of Edinburgh
ICMB
2004

Declaration

I declare that this is my own work. Any contributions made by others are clearly cited.

Contents

Chapter 3. Protein-protein interactions mediating the *fis* phenotype

Chapter 4. The requirement for *FIE* in the embryo and endosperm

Chapter 5. Discussion

References

Publications

I Acknowledgements

Firstly, I would like to thank the BBSRC for the funding of my PhD. I would also like to express thanks to my supervisor Dr Justin Goodrich for all his support, help, and patience throughout my time in the lab. Many thanks also to Dr Gwyneth Ingram and Dr Frédéric Berger for their valuable assistance with the microscopy. I am also very grateful to all members of the Goodrich lab, past and present, and in particular, Dr Colin MacDougall for taking me under his wing when I first joined the lab. I would also like to show particular gratitude to Tony Bishopp and Dr Daniel Schubert for their critical reading of some of the chapters.

In addition, I would like to thank my parents for their moral and financial support, without which I could not have completed this study. Last but not least, I would like to thank all my friends for putting up with me moaning all these years, especially Dr Lesley Gerrard who in addition to fuelling me with food and wine also took time out to read my thesis and spot (most of) my many grammatical and spelling mistakes.

II Abbreviations

BSA	bovine serum albumin
bp	base pair
cDNA	complementary deoxyribonucleic acid
dATP	2'- deoxyadenosine 5'- triphosphate
dCTP	2'- deoxycytidine 5'- triphosphate
dGTP	2'- deoxyguanosine 5'- triphosphate
dH_2O	sterile distilled water
DNA	deoxyribonucleic acid
dTTP	2'- deoxythymidine 5'- triphosphate
EDTA	ethylene-diaminetetraacetic acid
ITPG	isopropyl ß-D-thiogalactoside
Kb	kilobase
M	molar
mM	millimolar
mRNA	messenger RNA
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanogram
nm	nanometre
OD	optical density
PCR	polymerase chain reaction
PEG	polyethylene glycol
pmole	picomole
RNA	ribonucleic acid

rpm	revolution per minute
SDS	sodium dodecyl-sulphate
SSC	standard saline citrate
TE	Tris-HCL, ethylenediaminetetraacetic acid
Tris	tris(hydroxymethyl)aminomethane
Tween	polyethylenesorbitan monolaurate
μg	microgram
μl	microlitre
UTR	untranslated region
UV	ultra violet
X-Gal	4-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

III Abstract

In plants and animals, there are many examples where cell identity is maintained by epigenetic marks that prevent changes in cell type specific transcription programs, for example, in the control of homeotic gene expression patterns that specify regional fate during development (Goodrich and Tweedie 2002 and Orlando 2003). Central to the cell memory machinery are the Polycomb-Group (Pc-G) genes. The *Arabidopsis* (Pc-G) genes *FERTILISATION-INDEPENDENT ENDOSPERM (FIE), FERTILISATION INDEPENDENT SEED2 (FIS2)* and *MEDEA (MEA)* are required for ovule and seed development. These three genes confer similar mutant phenotypes in both unfertilized embryo sacs and developing seeds; that is, fertilization-independent endosperm proliferation and the abortion of fertilized seeds that carry a maternally inherited mutant *mea, fis2,* or *fie* allele.

I have shown that MEA and FIE, and MEA and FIS2 interact in yeast two-hybrid, and that the interaction of MEA and FIE is mediated by the amino-terminal region of MEA, whilst interaction of MEA and FIS2 requires the FIS2 C–terminal VEFS box. As the *FIE* and *MEA* RNA expression patterns over-lap *in vivo* and the proteins co-localize within the nucleus *in planta,* I conclude that the similarity of the *fie*, *fis2*, and *mea* phenotypes reflects the interaction of their gene products as part of a multiprotein complex. Furthermore, FIE, MEA, and FIS2 do not possess specific DNA binding activity, and it remains unclear how the FIE complex is directed to its target genes. Yeast two-hybrid assays were used to identify a novel FIE interacting protein: RAP-1, a member of the basic helix-loop-helix (bHLH) family of transcription factors. The biological relevance of this interaction is unclear, as it was not possible to confirm genetic interaction of *FIE* and *RAP-1 in vivo*. However, this could be due to redundancy as *RAP-1* is a member of a large gene family.

FIE is expressed in both embryo and endosperm, but predominantly affects endosperm development. In order to establish if *FIE* is needed in both, I used a strategy for targeted expression of *FIE* to the embryo and endosperm in a *fie* background. A rescue of the *fie* embryo lethality was achieved and viable *fie* homozygous seedlings were observed. This confirms that the transactivation strategy is potentially valuable for dissecting *fie* function. However, it was not possible to resolve the site of *fie* function due to technical problems with one of the transgenic lines used.

Chapter 1. Introduction

Chapter 1. Introduction

1.1 General introduction: plant development

A particular feature of plant development is that all adult structures are derived from meristems, groups of stem cells produced during embryogenesis. The entire aerial structure of plants develops from the SAM (shoot apical meristem), whilst the entire root system is derived from the cells that make up the root meristem. These meristems retain the capacity for cell proliferation throughout the adult life of the plant and can differentiate into many cell types. For instance, the SAM produces lateral organs (such as leaves), stem tissues, secondary meristems, as well as regenerating itself. The activity of meristems is re-iterative, for example, a vegetative meristem will repeatedly produce modular units each consisting of a bud, leaf, and internode. After a phase of vegetative growth, the SAM switches to an inflorescence meristem, which in turn can produce many floral meristems. Each floral meristem will produce the various floral organs, such as sepals, petals, stamens and carpels.

Unlike in animal development, plant cells divide in orientations that limit their movement between cell layers; also, the cell wall ensures that plant cells cannot migrate during development. As plant cells do not migrate they will not come into contact with many other cell types during development. It is therefore thought that clonally heritable restrictions on cell fate will be less important in plants compared with animals, and that positional information will play a more important role in plant cell fate. However, there is some evidence that heritable cell fate specification may be important during later stages of plant development (Carpenter and Coen 1990, and Jenik and Irish 2001). These restrictions in gene expression are determined by epigenetic changes, which in plants may be more easily reset than in animals. Thus, plant cells are totipotent and under certain environmental stimuli an adult somatic plant cell can have its cell fate memory completely erased and

consequently regain an ability to display its full genetic program as a somatic embryo, leading to the formation of a new plant. Accordingly, cloning is relatively easy in plants (Steeves and Sussex 1989). This is in contrast to animals, where epigenetic marks are much more stable, and are for the most part reset during embryogenesis and germline formation only. The term epigenetic is used in the modern sense here, where epigenetic changes confer mitotically and occasionally meiotically heritable changes in gene expression, and do not involve changes in the DNA sequence itself. In animals, epigenetic control of cell fate is mediated by two groups of genes: the trithorax-group (trx-G), which act to maintain expression of a specific set of developmental genes in those cells where they have been turned on, and the Polycomb-group (Pc-G), which act to maintain transcriptional repression in subsequent cell lineages, where the target genes have been initially turned off (Pirrotta 1998). The Pc-G consists of a structurally disparate group of proteins that act in large multiprotein complexes. Remarkably, several Pc-G proteins have been found to be conserved in plants, and regulate key developmental pathways such as flowering and seed development, although it is still unclear if their mechanism of action is also conserved.

To further characterise Pc-G gene action in plants, the aim of this thesis was to dissect the interactions and site of action of a group of Pc-G genes that collectively regulate seed development in the flowering plant model species *Arabidopsis thaliana*.

1.2 Angiosperm sexual reproduction

1.21 *The gametophyte*

One of the ways in that the plant life cycle differs from that of animals is that it alternates between two distinct multicellular phases: a haploid generation (the gametophyte), and a diploid generation (the sporophyte). In Angiosperms such as *Arabidopsis* the sporophyte is

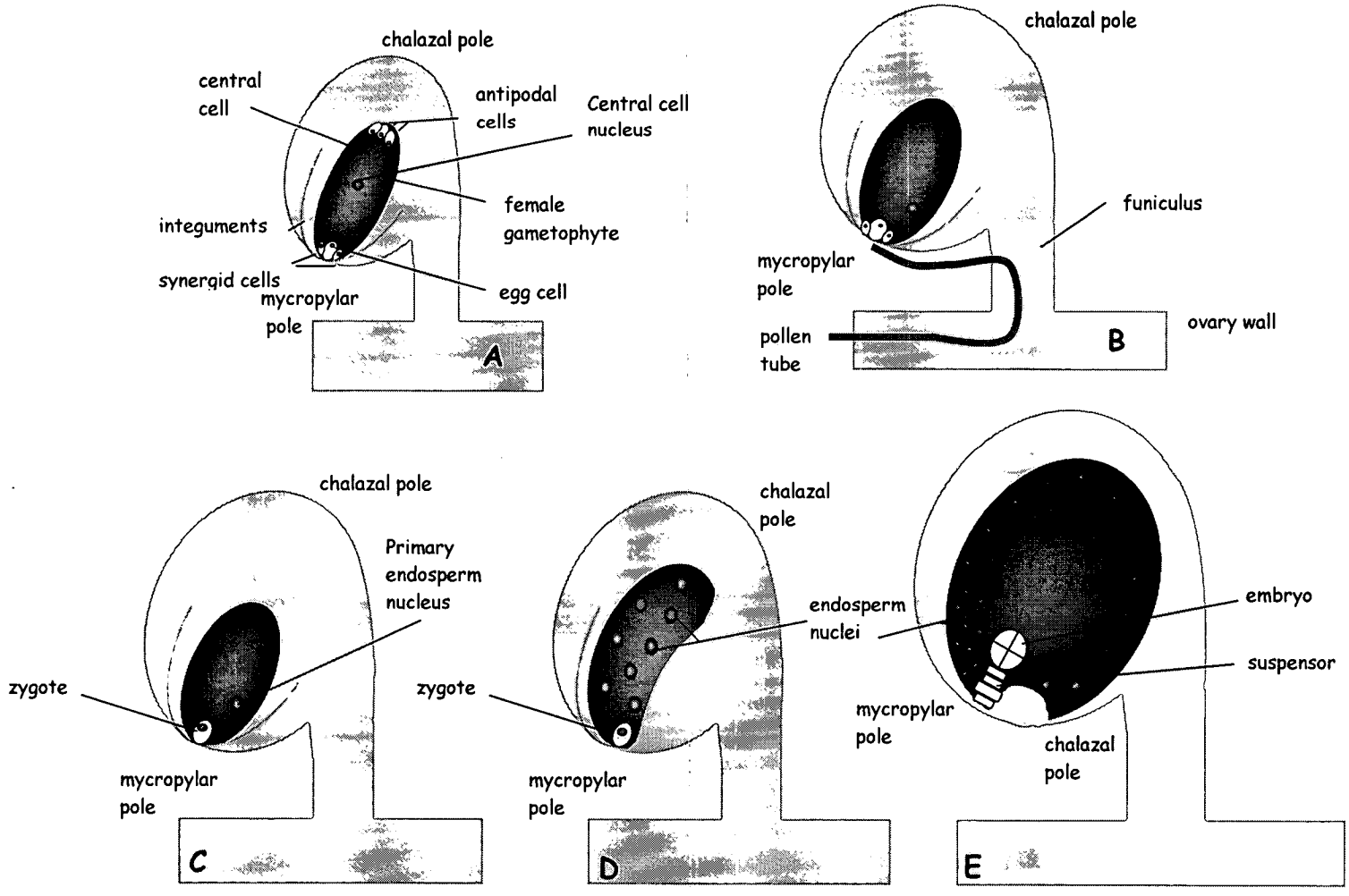


Figure 1.1 Schematic representation of an ovule and female gametophyte, before and after fertilization. Nuclei are represented by black circles. A) Unfertilized ovule. B) Ovule undergoing fertilization. C-E) Post-fertilization early seed development.

the dominant phase; this phase begins after fertilization with the zygote, and gives rise to the mature plant with vegetative organs and flowers that contain the reproductive organs. The gametophytic generation begins when haploid spores produced during meiosis undergo mitotic cell proliferation and differentiation to develop into the male and female gametophytes. The gametophytes consist of a small number of cells that are embedded within the male and female parts of the flower. The male gametophyte (or pollen grain), which develops within the stamen's anther, is a three-celled structure comprised of two sperm cells encased within a vegetative cell. The female gametophyte (embryo sac) develops within the ovules, which in turn are contained within the ovaries of the carpel. The female gametophyte typically consists of seven cells and four different cell types; three antipodal cells, two synergid cells, one egg cell, and one central cell that contains the two fused polar nuclei and occupies most of the volume of the embryo sac (figure 1.1A). The female gametophyte has a distinct polarity, with the egg and synergid cells at the micropylar pole, and the antipodal cells at the chalazal pole (figure 1.1A). Female gametophytic development requires gene products that are expressed both within the gametophyte itself and the surrounding sporophytic tissue (Drews *et al.* 1998, Grossniklaus and Schneitz 1998, and Schneitz 1999). In *Arabidopsis* the three antipodal cells degenerate to give the mature ovule (figure 1.1B). At the anterior pole the integuments form the micropylar pore, and at posterior or chalazal pole, the ovule joins the funiculus.

1.2.2 *Fertilization*

When pollen is transferred from anther to stigma the male gametophyte germinates a pollen tube that delivers the two sperm cells to the female gametophyte *via* the micropyle to one of the synergid cells (figure 1.1B) (Rotman *et al.* 2003). The synergid cell degenerates at the time of pollen tube discharge or very shortly before it (Faure *et al.* 2002). One of the sperm nuclei fuses with the egg cell to produce the diploid zygote, which develops into the embryo proper and the suspensor (an organ that is responsible for anchoring the embryo to the

embryo sac and ovule tissue). The other fertilizes the binucleate central cell to produce the triploid primary endosperm nucleus, which gives rise to the triploid endosperm (a zygotic body that supports embryo growth and development) (figure 1.1C-E). After this double fertilization event the integuments (the outer coating of the ovule) develop into the seed coat and the carpel elongates to form a fruit, which in *Arabidopsis* is termed the silique. Seed formation requires the coordinate development and cellular interactions of the embryo, the endosperm, and the integuments.

1.2.3 *Endosperm development*

Shortly after fertilization, the triploid primary endosperm nucleus undergoes successive divisions without cytokinesis. The resulting nuclei migrate into the expanding central cell producing a syncytium (or coenocyte) of a few hundred nuclei (Berger 1999). Each nucleus is surrounded by a denser cytoplasm, which together with the nucleus is known as the nuclear cytoplasmic domain (NCD) (figure 1.2). After three initial synchronous cycles of nuclear division three independent mitotic domains are observed in which nuclei either undergo coordinated divisions or increase in size. The three distinct mitotic compartments are organised along the anterior-posterior (A-P) axis. At the posterior pole, the nuclei are larger, and mitotic division does not occur, suggesting that ploidy in this region is greater than in the other mitotic domains (Boisnard-Lorig *et al.* 2001). By the time the embryo has reached its globular stage in development (figure 1.3), three morphologically distinct zones can be identified: chalazal endosperm (CZE) at the posterior pole, consisting of a cyst and several smaller nodules (made up of nuclei sharing a common NCD), peripheral endosperm (PE), and the dense micropylar endosperm (MCE) surrounding the embryo at the anterior pole (figure 1.2). It is thought that the endosperm is required for proper embryo formation, for example, the subtilisin-like serine protease ABNORMAL LEAF SHAPE (ALE1) is expressed strongly in the MCE throughout embryo development, and weakly in the embryo

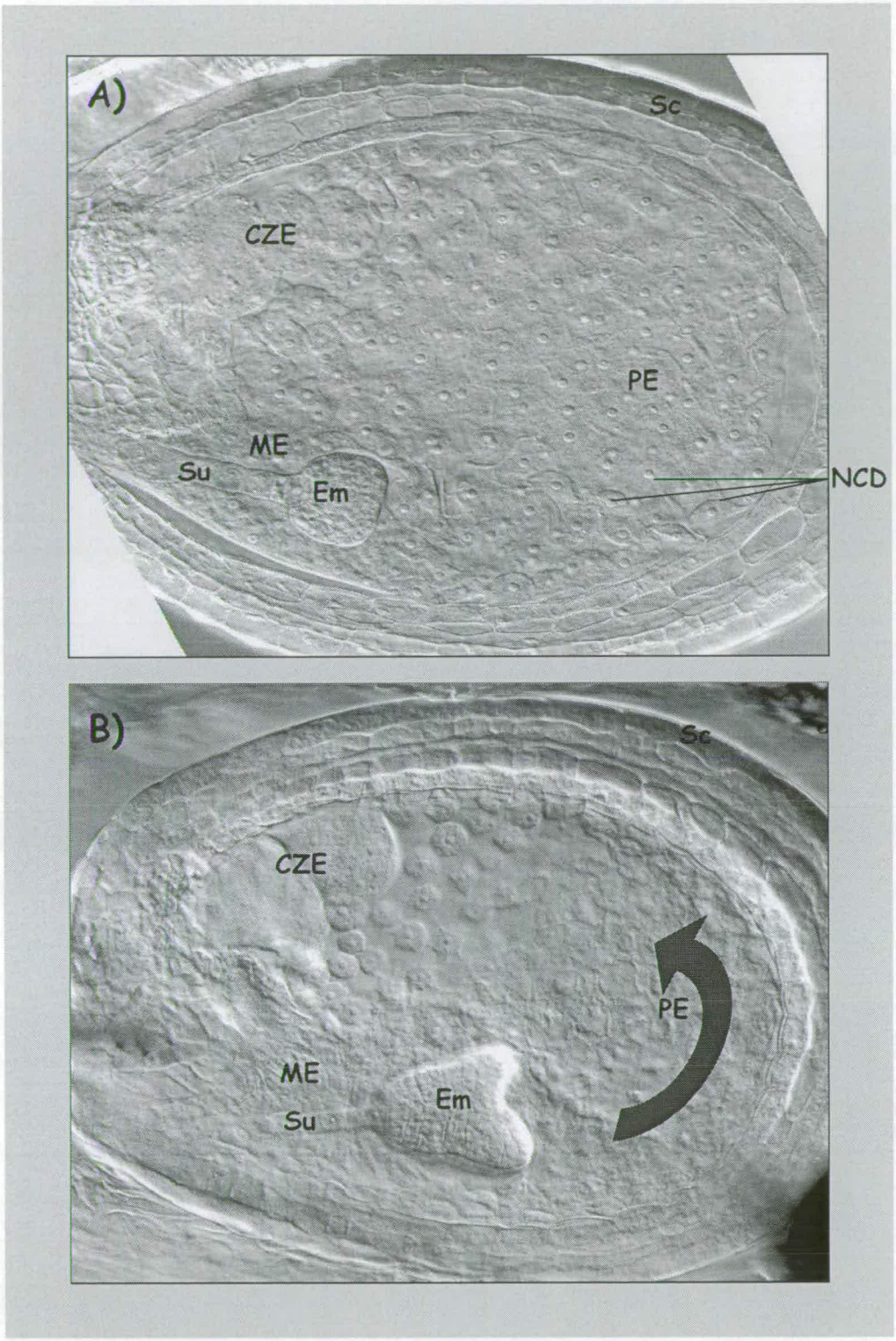


Figure 1. 2 Wild type developing seed. A) Transition-stage embryo before endosperm cellularization. B) Heart-stage embryo during endosperm cellularization. Cellularization occurs in a wave from the micropylar pole to the chalazal pole (anterior to posterior). The CZE remains uncellularized. CZE, chalazal endosperm, Em, embryo, ME, mycropylar endosperm, SC, seed coat, Su, suspensor, PE, peripheral endosperm, NCD, nuclear cytoplasmic domain. Arrow indicates direction of cellularization.

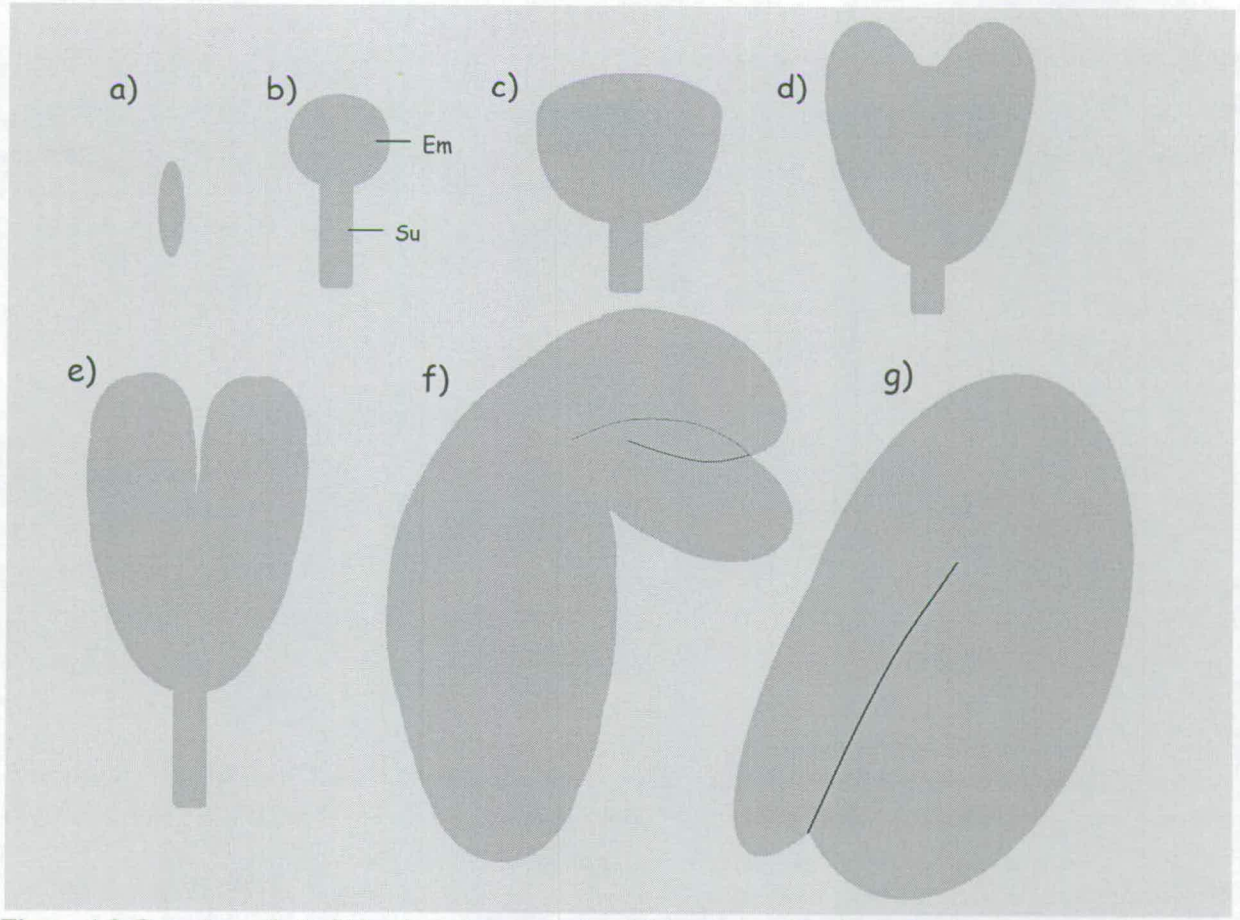


Figure 1.3 Overview of *Arabidopsis* embryogenesis. Main stages of embryo development. a) Zygote. b) Globular (preceded by the proembryo and dermatogen stages). c) Transition. d) Heart. e) Torpedo. f) Walking-stick. g) Mature. The suspensor senesces during the torpedo-stage. Em = embryo, Su = suspensor.

up until the late globular stage; where it affects the formation of cuticle at the interface between the embryo and the endosperm. When *ALE1* is mutated the cuticle is defective and the remnants of the endosperm remain attached to the embryo and the juvenile plants suffer excessive water loss and organ fusion (Tanaka *et al.* 2001).

Cellularisation of the endosperm is a gradual process that occurs shortly after the embryo globular stage, and is coupled with the eighth cycle of synchronous mitosis in the PE (Sørensen *et al.* 2002). Cellularisation begins in the area surrounding the embryo at the micropylar pole and progresses throughout the PE in the direction of the chalazal pole (figure 1.2B). The chalazal cyst and nodules remain uncellularised. Ultimately, in *Arabidopsis*, the endosperm is mostly degraded and absorbed by the developing embryo. This is in contrast to the cereals such as maize, rice, and barley where endosperm is much more persistent and makes up most of the bulk of the mature seed. The cereal endosperm contains stores of carbohydrates and proteins important in the feeding of humans and livestock, and represents around 60% of the world's food supply (Berger 2003). The study of endosperm development is therefore of high agronomic importance, and is the focus of much research.

1.2.4 *Embryo development*

The development of the fertilized egg cell into a mature embryo begins with an asymmetric cleavage of the zygote. The more apical of the resulting cells forms the embryo proper, and the other forms the suspensor. During development of the embryo an apical-basal (A-B) pattern is established consisting of: the SAM, the cotyledons, the hypocotyl, the root, and the root apical meristem. Along its radial axis the embryo comprises three main tissue types: the outer protoderm (epidermis), the middle ground meristem, and the inner procambium (vascular tissue) (figure 1.4). The A-B axis is established early on in development, and is

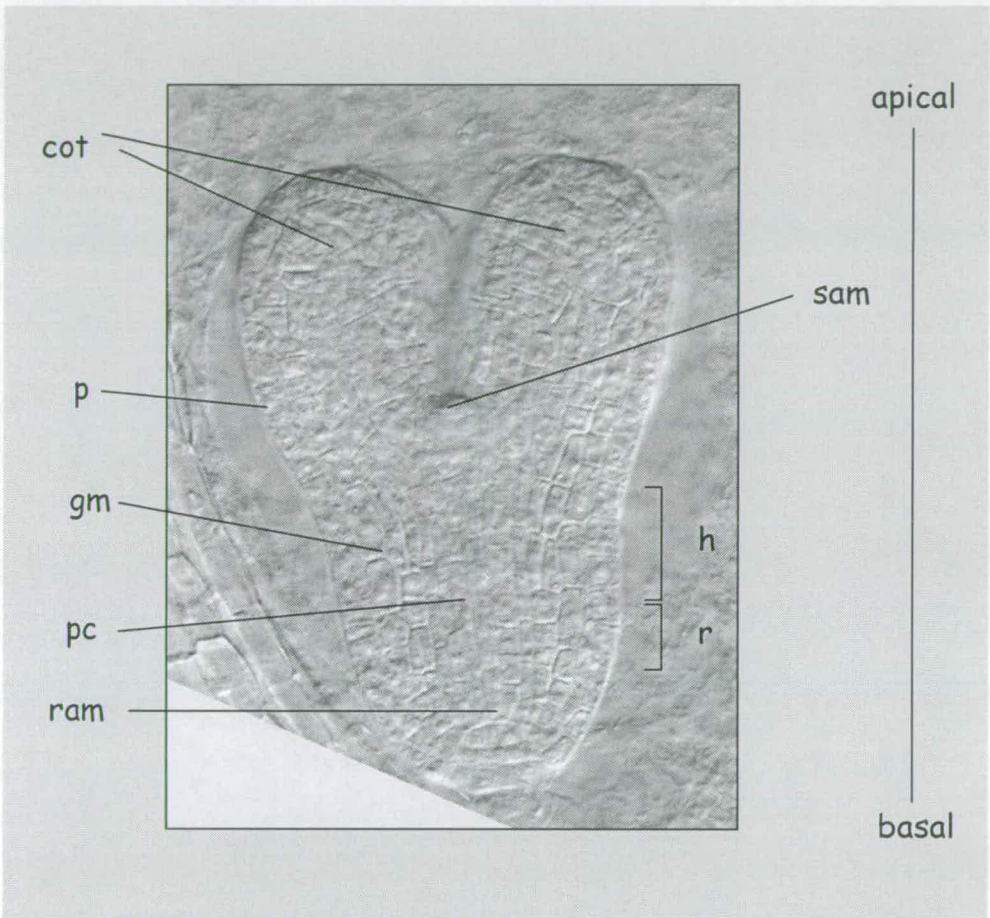


Figure 1.4 Embryo patterning. Torpedo-stage embryo, showing the three major embryonic tissue systems: protoderm (p), ground meristem (gm), and procambium (pc), and the five major organs/structures along the apical-basal axis: cotyledons (cot), shoot apical meristem (sam), hypocotyl (h), root (r), and root apical meristem (ram). DIC light microscopy.

thought to be maternally influenced, for example, *short integument1* (*sin1*) mutant mother plants show defects in the A-B axis regardless of the paternal gene input (Ray *et al.* 1996).

The morphology of the embryo changes throughout embryogenesis and is summarised in figure 1.3. The endosperm is also thought to play a role in embryogenesis, for example in controlling embryo size (Hong *et al.* 1996 and Garcia *et al.* 2003). Embryogenesis can occur without apparent surrounding maternal tissue or endosperm, for example, somatic cells can undergo embryogenesis in culture to produce viable plants (Williams 1986 and Dejong 1993), and egg cells fertilized *in vitro* can give rise to fertile adult plants (Kranz and Dresselhaus 1996). However, it is thought that the cell cultures secrete many different proteins into the medium, a process that contributes to the conditioning of the medium allowing it to mimic the nursing function of the endosperm (van Hengel *et al.* 1998). Overexpression of certain genes can also result in somatic embryogenesis, for example on the leaves of plants ectopically expressing *LEAFY COTYLEDON 1* (*LEC1*) (Lotan *et al.* 1998).

Embryogenesis ends with a dormancy period; after the embryo has reached its maximum size and large amounts of storage products have been synthesised the seed becomes dehydrated and metabolic activities cease. Dormancy is maintained by high abscisic acid (ABA) levels, which prevent the seed from germinating whilst still within the silique (Koorneef *et al.* 1984 and Nambara *et al.* 1992). When environmental conditions are suitable, a seed is capable of undergoing germination. Germination is a complicated process including rehydration, repair mechanisms, cell elongation and the expression of the extensin-like protein AtEPR1 to aid the emerging radicle (Dubreucq *et al.* 2000).

1.3 Apomixis

Certain plants have the ability to produce seeds asexually. In this process, known as apomixis, female gametes develop without meiosis (or with abnormal meiosis) and embryos (and sometimes endosperm) develop without fertilization (Koltunow 1995). Although it is a complex process, apomixis is often inherited as a simple Mendelian trait, which may suggest that it is controlled by relatively few genes (Kindiger 1996). It is thought that apomixis may have evolved through modifications of the normal sexual reproduction pathway, rather than constituting a novel pathway distinct from sexual reproduction (Grimanelli *et al.* 2001). *Arabidopsis thaliana* is a non-apomictic plant and seed development requires a double fertilization event.

The ability to develop new strains of apomictic crop plants has important implications for agriculture, one application being hybrid seed production. Traits transferred into an apomictic plant, whether by classical breeding or by genetic engineering, would not be lost in the genetic shuffling that occurs during sexual reproduction. In crops, the production of hybrids is complex and expensive; apomixis could therefore reduce the cost of seed production, while ensuring the seed retain the benefits of hybrid vigour, essentially allowing clonal propagation through seed (Grimanelli *et al.* 2001). It could also avoid the degeneration of breeding stocks of some vegetatively propagated plants that accumulate pathogens through repeated use (Moffat 2001).

1.4 Fertilization-independent seed development

1.4.1 *Identification of mutants causing fertilization-independent seed development*

A screen designed by Ohad *et al.* (1996) used a conditional male sterile line to identify mutant *Arabidopsis* plants that undergo any reproductive processes in the absence of

fertilization. This led to the isolation of the *FIE* (*FERTILIZATION-INDEPENDENT ENDOSPERM DEVELOPMENT*) gene (Ohad *et al.* 1996). In the absence of fertilization, wild type plants show no elongation of the carpels, and no seed development (figure 1.5A), whilst *fie* mutant plants exhibit carpel elongation, and some seed like structures are seen displaying seed coat development. These seed-like structures show autonomous endosperm development. The central cell replicates (without fertilization), allowing endosperm development to progress until the syncytium stage (figure 1.5B). The egg cells in these mutants do not undergo replication prior to fertilization. Close to 50% of the female gametophytes in heterozygous plants show some limited endosperm development; this 1:1 segregation suggests that the *fie* mutation is gametophytic in nature (Ohad *et al.* 1996).

Chaudhury *et al.* (1997) also found similar phenotypes in another genetic screen using a male sterile background to isolate mutants displaying aspects of seed development without fertilization. Three fertilization-independent seed mutants, designated *fertilization-independent seed* (*fis*), were identified and shown to define three separate loci: *FIS1*, *FIS2*, and *FIS3* (Chaudhury *et al.* 1997). It has since been shown by sequence analysis and complementation that *FIS3* and *FIE* are allelic. Grossniklaus *et al.* (1998) independently identified a gametophytic maternal effect gene named *MEDEA*[*] (*MEA*), which is allelic to *FIS1*. In *mea* and *fis2* however, autonomous endosperm development progresses to the cellularised stage, and occasionally zygote-like and early embryo-like structures form (Chaudhury *et al.* 1997, Grossniklaus *et al.* 1998).

The fact that *fis* mutants have normal meiosis and embryo sac formation indicates that the *FIS* genes are not likely to be the initial controlling genes for apomictic seed development, but they are likely to be inactivated so that full development of apomictic seeds can proceed (Chaudhury *et al.* 1997).

[*] Euripides (431 B.C.), *Medea.* Medea kills her two children in revenge for Jason's infidelity.

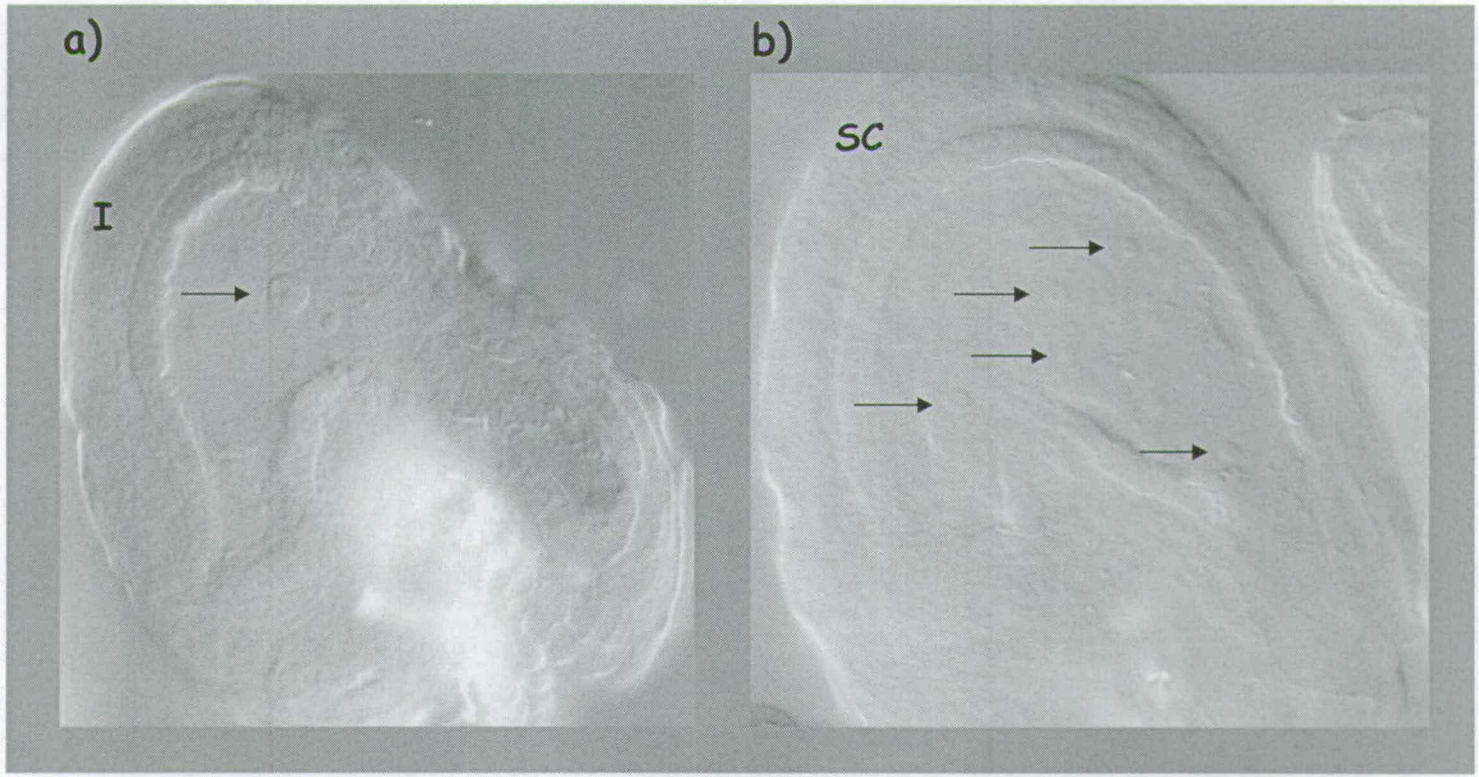


Figure 1.5 Autonomous endosperm development in *fie* mutant ovules. a) Wild type unfertilized ovule. Arrow indicates central cell nucleus. b) Unfertilized *fie* ovule displaying some aspects of seed development. Arrows indicate endosperm nuclei. Integuments, I, Seed Coat, SC.

1.4.2 Maternal effects of *FIS* mutants on seed development

When fertilization is not prevented and mothers heterozygous for a *fis* mutation are self-pollinated or pollinated with *FIS*+ pollen, 50% of the resulting seeds are viable, but the remainder contain an embryo that arrests at the globular to torpedo stages of development and subsequently collapses during desiccation (figure 1.6). Occasionally in *mea* and *fis2* mutant seed, the embryo germinates to produce plants generating normal rosette leaves and an inflorescence that produce siliques containing nearly 100% aborting seed (Kiyosue *et al.* 1999 and Chaudhury *et al.* 1997).

In all three *fis* mutant seed the endosperm is abnormal. In these mutants the endosperm shows signs of overproliferation of the endosperm; the CZE domain in particular is enlarged and expands towards the anterior pole, resulting in ectopic nodules and cysts (Sørensen *et al.* 2001) (figure 1.7). The exception to this is the *mea-1* and *mea-2* alleles, which have been reported to produce giant heart-stage embryos and endosperm with fewer nuclei than wild type (Grossniklaus *et al.* 1998). Because immature *mea* mutant embryos can be rescued from seeds before they abort, and cultured *in vitro* to give rise to phenotypically normal, but female sterile plants, the *mea* phenotype may in part result from defects in the endosperm, *i.e.* the effects on the embryo could be indirect and due to the surrounding defective endosperm. The *fis* mutants also display ectopic expression of the posterior marker gene linked to the enhancer trap line KS117, suggesting that the *FIS* genes may be involved in the A-B patterning of the endosperm (Sørensen *et al.* 2001). Presumably the *FIS* genes act to repress endosperm development before and after fertilization, and may also be required in the embryo.

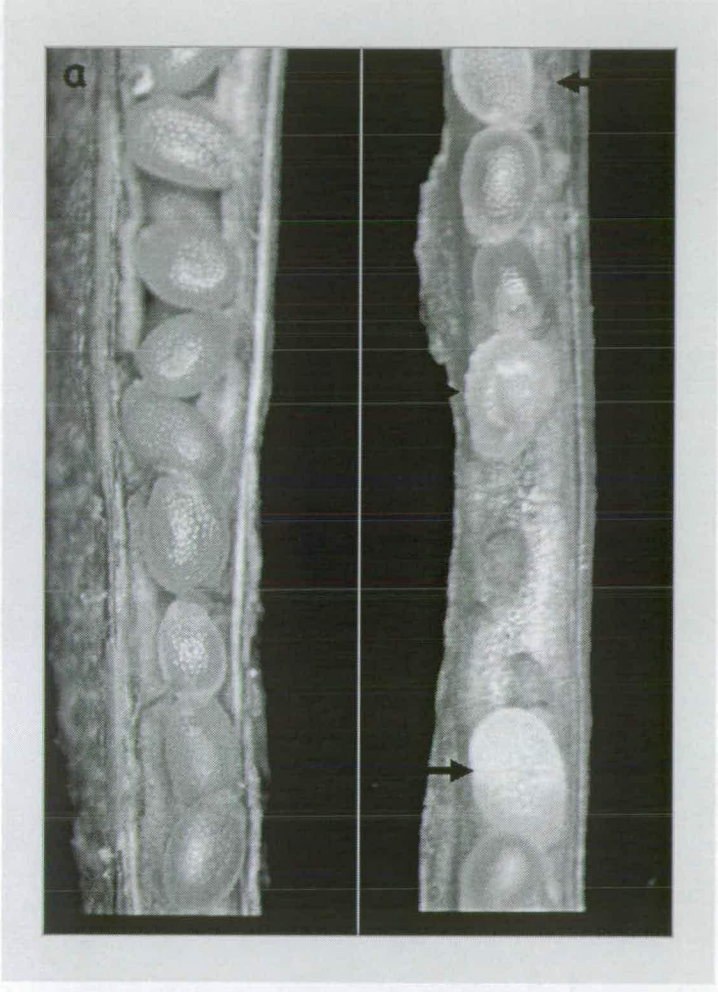


Figure 1.6 Maternal effect of *fie* mutation. a) Wild type dissected silique. b) *fie* heterozygote dissected silique. Arrows indicate aborting seed.

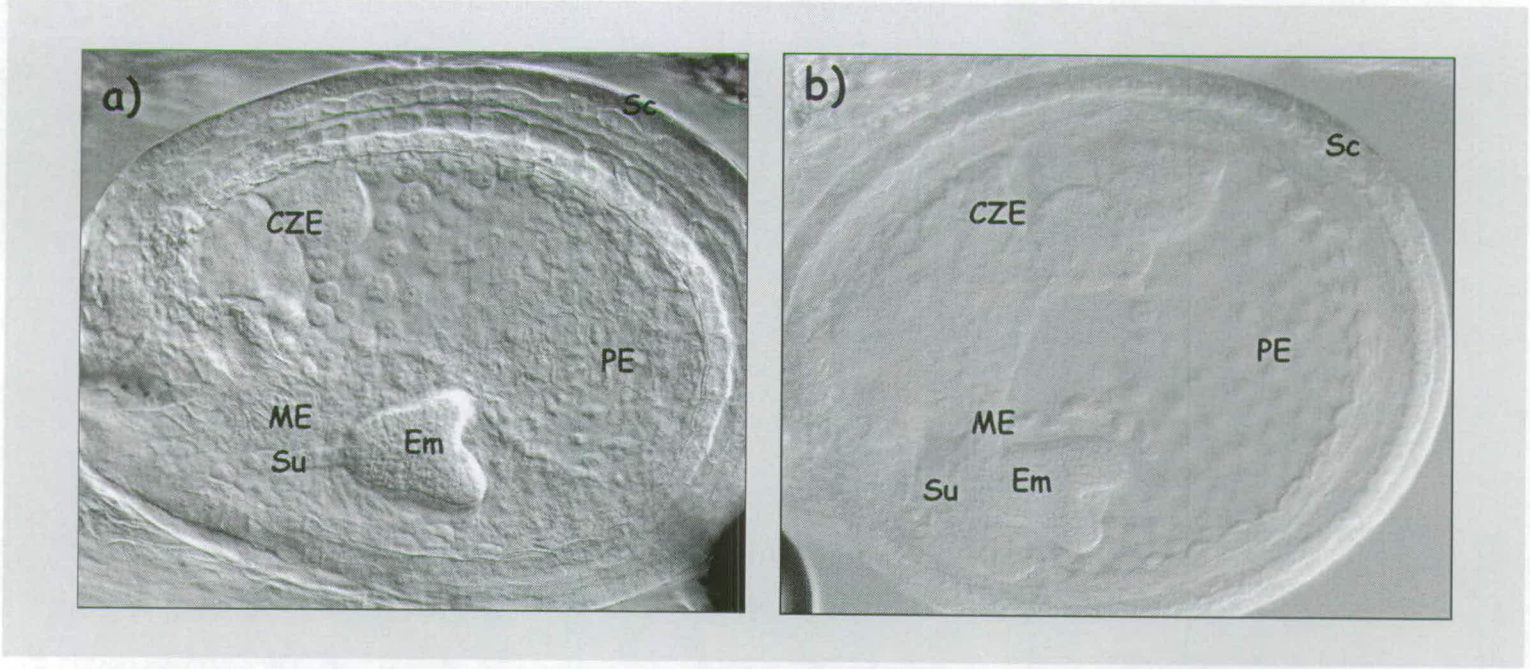


Figure 1.7 Effects of *fie* on seed development. a) Wild type developing seed at heart-stage of embryo development. b) *fie* seed of similar stage. ME and PE remain uncellularized, CZE is abundant and expansive, embryo is abnormal and aborts. CZE, chalazal endosperm, Em, embryo, ME, mycropylar endosperm, SC, seed coat, Su, suspensor, PE, peripheral endosperm.

1.4.3 Spatial pattern of *FIS* gene expression during ovule and seed development

The *fis* mutations affect the gametophyte, endosperm and embryo, suggesting that they might be expressed in some or all of these tissues. The *FIS* expression patterns have been investigated using several different approaches: directly, by *in situ* localisation of mRNA transcripts, and indirectly, using various reporter constructs in transgenic plants. The pattern for each gene tends to vary with each different method; this in part could be due to the different ecotypes used in each experiment, or that different reporters do not faithfully replicate endogenous gene expression. Luo *et al.* (2000) fused the *FIS* gene promoters and coding sequences to the glucuronidase (*GUS)* reporter gene and transformed wild type plants. Their results showed that the *p(FIS2)::FIS2-GUS and p(MEA)::MEA-GUS* products are found in the embryo sac, in each of the polar cell nuclei, and in the central cell nucleus. After fertilization the maternally derived protein occurs in the nuclei of the syncitial endosperm. Before cellularisation, expression terminates in the MCE and PE, and later in the CZE (Luo *et al.* 2000). *FIE* expression patterns were shown to overlap with that of *FIS2* and *MEA*, being expressed in the female gametophyte before fertilization, and in the endosperm after fertilization. *p(FIE)::GUS* and *p(FIE)::GFP* were also expressed in the embryo at the torpedo, walking stick, and early maturation stages (Yadegari *et al.* 2000). *p(FIE)::GUS* and *p(FIE)::GFP* however, were found to be expressed in all endosperm domains in a biphasic mode, beginning in the central cell, downregulated during the late syncytial phase, and reactivated at the transition-stage of embryo development whereby they remain expressed throughout the rest of development (Yadegari *et al.* 2000). Presumably the early endosperm expression corresponds to maternal expression and the later endosperm and embryo expression corresponds to zygotic expression of the transgenes. *In situ* hybridisation data is fairly consistent with the reporter gene data but differs in that *MEA* and *FIE* mRNA are seen to be expressed in the embryo and suspensor until the heart and torpedo stages of

development (Vielle-Calzada *et al.* 1999 and Spillane *et al.* 2000). If the *p(FIS2)::FIS2-GUS* endosperm specific expression correctly reflects endogenous FIS2 expression, it would suggest that the embryo abortion phenotype seen in *fis2* mutants is an indirect consequence of endosperm defects. Unlike the other *FIS* genes, *FIE* is expressed beyond seed development in seedlings and mature plants and is therefore thought to be required for vegetative development (Ohad *et al.* 1999, Luo *et al.* 2000).

1.4.4 Imprinting may account for maternal effects of *FIS* genes

The wild type *FIS* genes exert their effect by maternal gametophytic control during seed development, *i.e.* a wild type maternal allele must be inherited for normal embryo and endosperm development irrespective of the paternal allele inherited. One explanation for this is that maternal expression of the *FIS* genes in the female gametophyte is required before fertilization for normal embryo/seed development post-fertilization. For example, maternal loading of *FIS* RNA or protein into the egg or central cell before fertilization that is required post-fertilization (as in the case of PROLIFERA (PRL), a MCM-related protein that is required for DNA replication during the S-phase of the cell cycle (Springer and Holding 2002)). However, an alternative hypothesis is that *FIS* genes are imprinted such that only the maternal allele is active during seed development. In this case, the paternal *FIS+* allele would be silent and therefore unable to rescue the *fis* maternal allele. This pattern of imprinting of the *FIS* genes would be consistent with the parental conflict hypothesis of Haig and Westoby (Haig and Westoby 1989 and 1991), which describes endosperm as playing an analogous role to the placenta in animal systems and concludes that genes that restrict zygote growth become paternally silenced, whereas genes promoting growth would be silent from maternal alleles (Scott *et al.* 1998a, Scott *et al.* 1998b Mora-Garcia *et al.* 2000). However, a complication is that it has been shown that there is a delay in paternal activation for a large

number of loci during the first few divisions after fertilization in both the embryo and endosperm (Vielle-Calzada *et al.* 2000). This paternal delay in activation is thought to be genome-wide, although there are exceptional loci with earlier activation (Vielle-Calzada *et al.* 2001 and Weijers *et al.* 2001).

There are several lines of evidence suggesting *MEA* is paternally imprinted for longer than the genome-wide phenomenon in the endosperm at least: using differential expression of maternal and paternal *MEA* alleles detected by polymorphisms between different ecotypes (Kinoshita *et al.* 1999), and using reporter constructs (Luo *et al.* 2000, Yadegari *et al.* 2000, Vielle-Calzada *et al.* 2000). Luo *et al.* (2000) reported paternal *MEA* expression in the chalazal cyst two days after pollination, whilst paternal *FIE* was expressed three days after pollination in the early heart embryo and in the cellularised endosperm. Kinoshita *et al.* (1999) however, found paternal *MEA* to be expressed only in the embryo, and is silent throughout endosperm development. Paternal *FIS2* however, was reported to be silent at all developmental stages in the developing seed, although even the maternal copy is silent at the later stages of seed development (Luo *et al.* 2000). Together these results suggest that prolonged paternal silencing of *MEA* could account for parent-of-origin effects as *MEA* is clearly imprinted in the endosperm. However, it is less clear whether *FIS2* and *FIE* are imprinted beyond the normal period at which zygotic or paternal transcription begins.

Furthermore, it has recently been shown that a mutation in the *ddm1* gene *(decrease in DNA methylation 1)* can rescue *mea* seeds; this was thought to be by activating the paternally inherited wild-type *MEA* allele. Rescued seed are larger than wild type seeds due to an enlarged embryo and a persistent large portion of partially cellularised endosperm surrounding the embryo. The enlarged seed also exhibit some of the abnormalities found in aborting *mea* seeds, such as delayed morphogenesis and abnormal cell proliferation,

suggesting that the rescue is probably incomplete. However, unlike *mea* seeds the enlarged seeds are able to resume embryogenesis and often complete their development. By contrast, *ddm1* does not rescue the maternal lethality of *fie* alleles (Vielle-Calzada *et al* 1999). All three *fis* mutant seeds however, can be rescued by pollen containing a hypomethylated genome and carrying an antisense *DNA methyltransferase-1* (*MET-1*) transgene (Luo *et al.* 2000, Adams *et al.* 2000). When *mea/mea* and *fis2/fis2* homozygous plants were pollinated with *MEA/mea* anti-*MET-1* or *FIS2/fis2* anti-*MET-1* pollen nearly 100% of seed were rescued compared to the 50% that would be expected if the rescue depended on reactivation of the paternal *MEA/FIS2* alleles. Luo *et al.* (2000) also demonstrated that the paternal p(*MEA*)::*MEA-GUS* and p(*FIS2*)::*FIS2-GUS* transgenes were not reactivated in anti-*MET-1* hypomethylated plants, again suggesting that rescue did not involve reactivation of paternal *FIS* alleles. This indicated that the rescue of *fie* and *fis2* endosperm by demethylation of the pollen genome results in the derepression of silenced genes other than the *FIS* genes which act as modifiers of the maternal effect of *fis* mutations (Luo *et al.* 2000).

Recent studies of the *DEMETER*[*] (*DME*) gene have given more insight into *FIS* imprinting. Mutants for *DME* show a maternal gametophytic effect and phenotype similar to that of *fis* mutants. *DME* is expressed in the central cell prior to fertilization and diminishes shortly after fertilization and encodes for a DNA glycosylase that is required for transcriptional activation of *MEA* (Choi *et al.* 2002). It is thought that *MEA* might be silenced by default in the female gametophyte by the DNA-modifying enzyme MET1 methyltransferase (Xiao *et al.* 2003) and becomes active in the central cell because of the action of *DME*, and that the restriction of *DME* expression to the female gametophyte prevents activation of paternal *MEA* allele transcription in the endosperm (Choi *et al.* 2002 and Dickinson and Scott 2002).

[*] Demeter of Cnidos, Goddess of Fertility and Mother of Corn.

DME does not appear to regulate *FIS2* and *FIE*, suggesting that either MEA is the only genuinely imprinted *FIS* gene, or that *FIE* and *FIS2* are imprinted by a different mechanism.

The *FIS* genes themselves may also regulate imprinting as paternal excess of the normal balance of maternal (m) / paternal (p) genomes (2m/1p) can result in phenotypes similar to that observed in *fis* seeds. Maternal excess results in precocious cellularisation of the endosperm and its reduced growth (Scott *et al.* 1998).

1.4.5 The *FIS* genes encode Polycomb-group proteins

1.4.5.1 *FIE*

FIS3/FIE is a single copy gene and shares homology (~40% amino-acid sequence identity) with the *Drosophila* Polycomb-Group (Pc-G) gene *EXTRA SEX COMBS (ESC)* (Ohad *et al.* 1999), which encodes a WD40 motif repeat protein. The WD domain typically contains a GH (glycine and histidine) dipeptide 11-24 residues from its N-terminus and a WD (tryptophan and aspartic acid) dipeptide at its C-terminus and is 40 residues long (hence the name WD40); between GH and WD lies a conserved core. The *FIE* sequence contains seven repeats of the WD40 motif (Ng *et al.* 1997) (figure 1.8A). The crystal structure of the model WD40 protein G_β (Sondek *et al.*1996 and Wall *et al.* 1995) suggests that FIE folds into a β-propeller structure (figure 1.8B), consisting of seven blades, each comprising a twisted, four-stranded β-sheet organised into a toroid (doughnut shape). Each WD40 motif contributes the inner 3 β-strands of one blade; the outer (4^{th}) strand is formed by more variable sequences following each WD motif. β-propeller structures have diverse functions beyond transcriptional regulation, but are thought to have the common function as protein assembly platforms.

a)

```
FIE MS--KITLGNE--------------------------------SIVGSL------TPSN---------------- 19
ESC MSSDKVKNGNEPEESEESCGDESASYTTNSTTSRSK-------SPSSSTRSKRR----------GRRST-KSKPK 57

FIE ----KKSYKVTNRIQEGKK-PLYAVVFNFLDARFFD-VFVTAGGNRITLYNCLGDGAISALQSYADEDKEESFYT 88
ESC SRAAYKYDTHV---KENHGANIFGVAFNTLLGKDEPQVFATAGSNRVTVYECPRQGGMQLLHCYADPDPDEVFYT 129
                  _________________WD1_________________          ____WD2___

FIE VSWACGVNGN-PYVAAGGVKGIIRVIDVNSETIHKSLVGHGDSVNEIRTQPLKPQLVITASKDESVRLWNVETGI 162
ESC CAWSYDLKTSSPLLAAAGYRGVIRVIDVEQNEAVGNYIGHGQAINELKFHPHKLQLLLSGSKDHAIRLWNIQSHV 204
    _____________________________          _____________________WD3_____________

FIE CILIFAGAGGHRYEVLSVDFHPSDIYRFASCGMDTTIKIWSM--KEFWTYVEKSFTWTDDPSK--FPTKFVQFPV 233
ESC CIAILGGVEGHRDEVLSIDFNMRGD-RIVSSGMDHSLKLWCLNTPEFHHKIELSNTFSQEKSTLPFPTVTKHFPD 278
            _______________WD4_______________

FIE F-TASIHTNYVDCNRWFGDFILSKSVDNEILLWEPQLKENS-----PGEGASDVLLRYPVPMCDIWFIKFSCDLH 302
ESC FSTRDIHRNYVDCVQWFGNFVLSKSCENAIVCWKPGQLHQSFEQVKPSDSSCTIIAEFEYDECEIWFVRFGFNPW 353
        _____________WD5_____________                          __________WD6___

FIE LSSVAIGNQEGKVYVWDLKSCPPVLI--TKLSHNQSKSVIRQTAMSVDGSTILACCEDGTIWRWDVI-TK      370
ESC QKVIALGNQQGKVYVWELDPSDPEGAHMTTLHNSRSVATVRQIAFSRDASVLVYVCDDATVWRWNRRQTTSI    425
    ___________________          _________________WD7_________________
```

b)

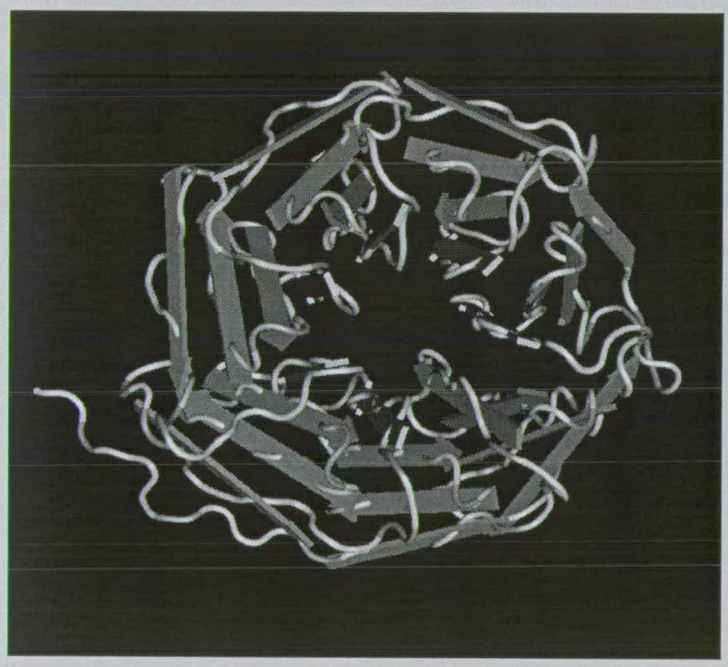


Figure 1.8 Predicted FIE protein structure. a) Primary sequence alignment between ESC and FIE. b) The 7-bladed beta-propeller structure of the model WD40 protein, G_{β}, FIE is predicted to fold similarly. Figure taken from NCBI conserved domain database viewed with Cn3D 4.1 (http://www.ncbi.nlm.nih.gov /Structure/cdd/cdd.shtml).

1.4.5.2 *MEA*

Cloning and sequencing of *MEA/FIS1* revealed structural homology to the *Drosophila* Pc-G protein *ENHANCER (OF ZESTE) (E(z))* (Kiyosue *et al.* 1999). There are three main regions of homology between MEA and E(Z): the C-terminal SET domain, which is named after the three founding members of the protein-family in *Drosophila: Suppressor of variegation 3-9 [Su(var)3-9], E(z),* and *Trithorax (Trx)*; N-terminal to the SET domain lie two cysteine-rich regions consisting of the CXC domain and five additional highly conserved cysteine residues (figure 1.9). A major breakthrough in the Pc-G field was the recent identification of a biochemical function for the SET domain. The SET domain, along with the adjacent cysteine-rich regions confers histone-specific methyltransferase activity (Rea *et al.* 2000, Czermin *et al.* 2002). Such modification of histone tails plays an important role in regulating higher-order chromatin structure and gene regulation.

The difference between the *mea* alleles in phenotype, discussed previously (in section 1.5) may be attributed to the fact that *mea-1* and *mea-2* products are predicted to contain important domains upstream of the SET domain, whilst the *mea-3* and *mea-4* products are predicted to lack most, if not all of the protein domains required for Polycomb activity (Kiyosue *et al.* 1999).

1.4.5.3 *FIS2*

FIS2 is predicted to encode a Zinc finger transcription factor containing many repeats of 17 and 22 amino acids and a VEFS-box (Luo *et al.* 1999). The VEFS-box was identified as a region of sequence similarity between the *Arabidopsis* proteins VERNALISATION (VRN2), EMBRYONIC FLOWER 2 (EMF2), and the *Drosophila* Pc-G protein SUPPRESSOR OF ZESTE 12 (Su(Z)12)(Birve *et al.* 2001) (figures 1.10 and 1.11). The biochemical function of the box is currently unknown.

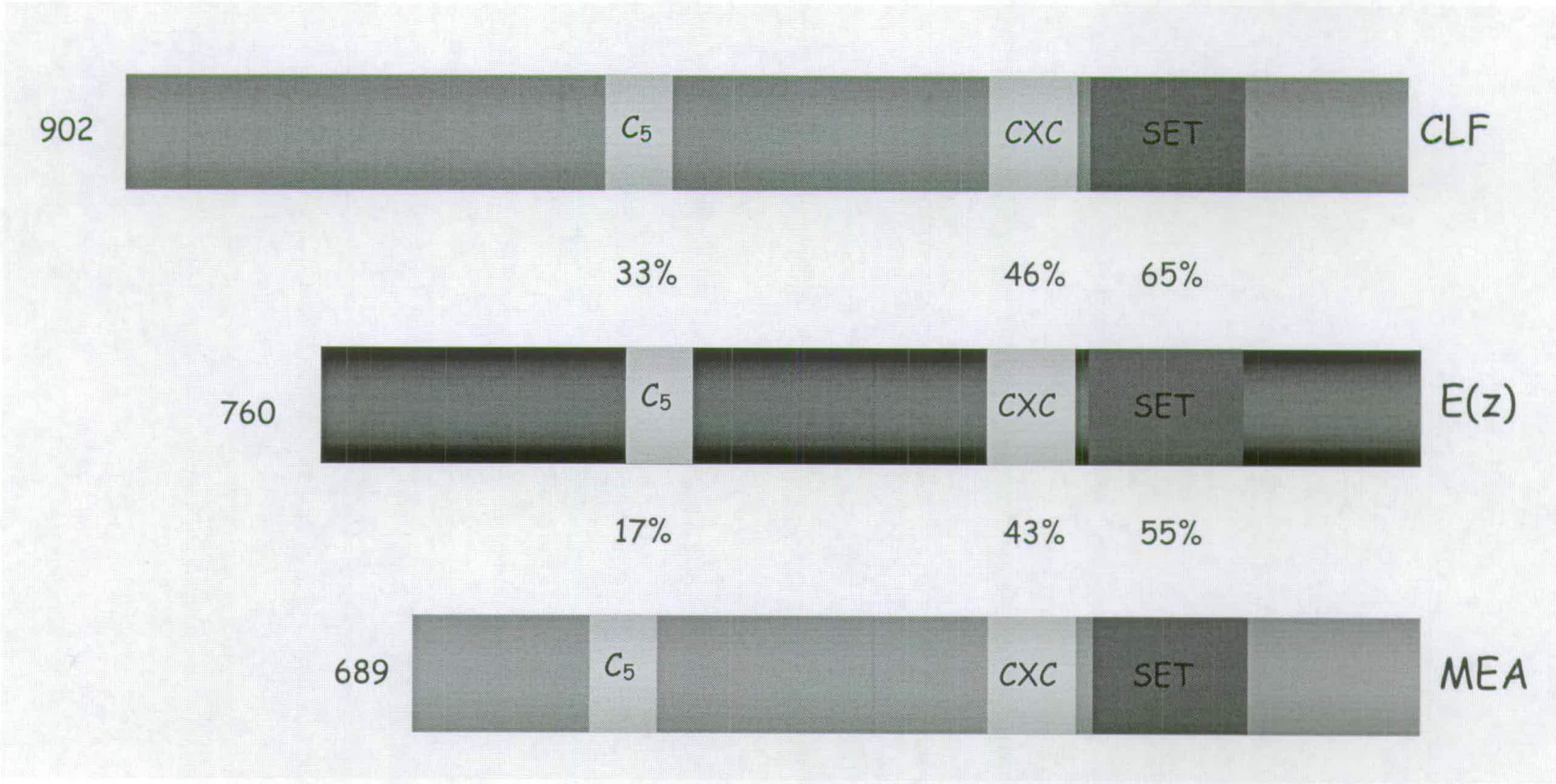


Figure 1.9 ***Drosophila*** **protein E(Z) showing regions of homology with the** ***Arabidopsis*** **homologues, CLF and MEA.** Not drawn to scale. There are three main conserved regions: the C-terminal SET domain, which is named after the three founding members of the protein family in *Drosophila: Suppressor of variegation 3-9 [Su(var)3-9], E(z),* and *Trithorax (Trx).* N-terminal to the SET domain lie two cysteine rich regions consisting of the CXC domain and five additional highly conserved cysteine residues (C_5). Numbers show percentage of amino-acid conservation. Protein lengths in amino acids are shown on the left.

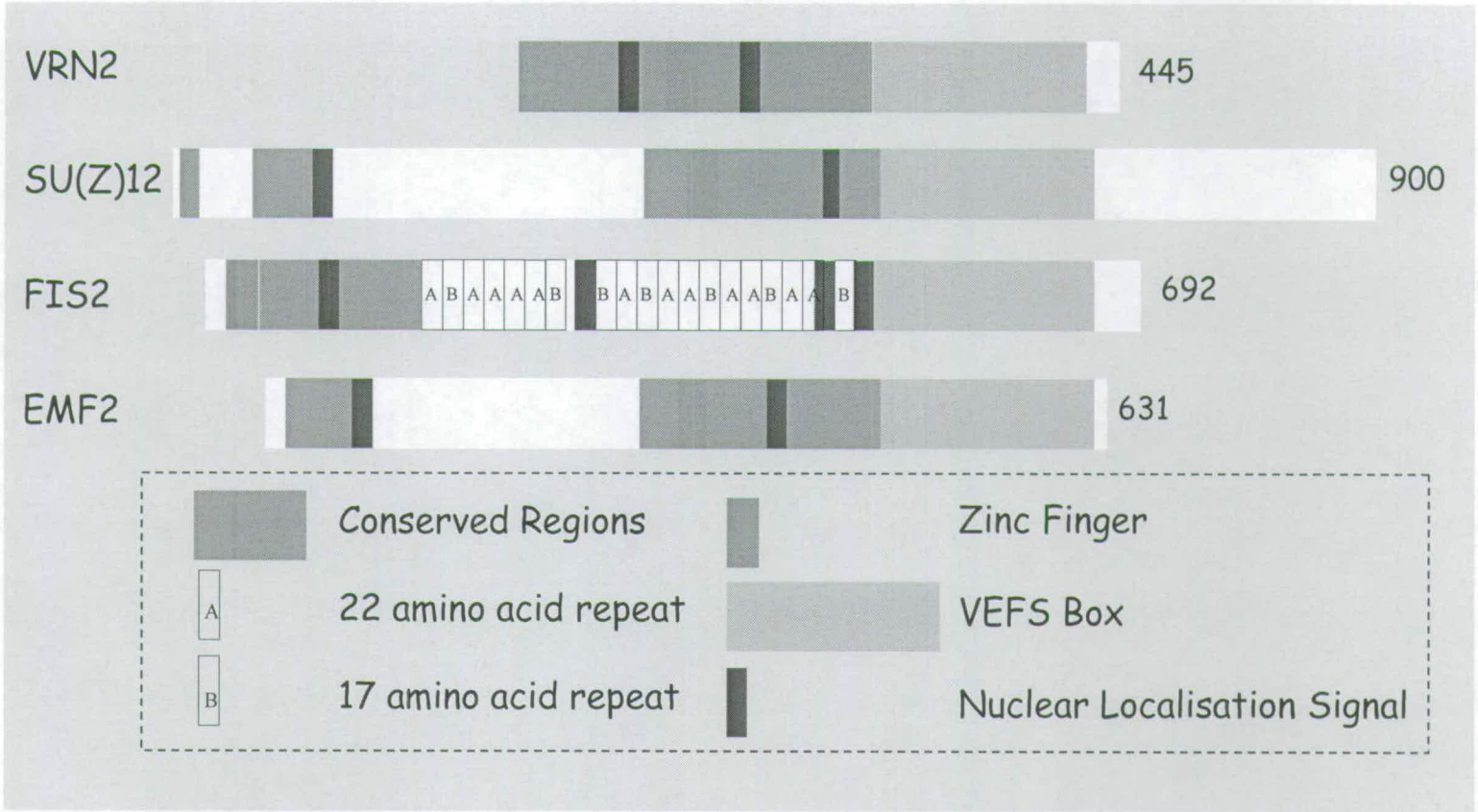


Figure 1.10 The FIS2 protein family. Adapted from Gendall *et al. CELL.* **107 (4):** 525-535 (2001). Protein lengths in amino acids are shown on the right. Regions showing similarity shaded in grey. The zinc finger motif is indicated in jade, and positions of putative NLSs/clusters of basic residues indicated in blue. The position of the VEFS box is shown in green. The 22 amino acid A and 17 amino acid B repeats of FIS2 are indicated. Regions with no significant similarity are unshaded

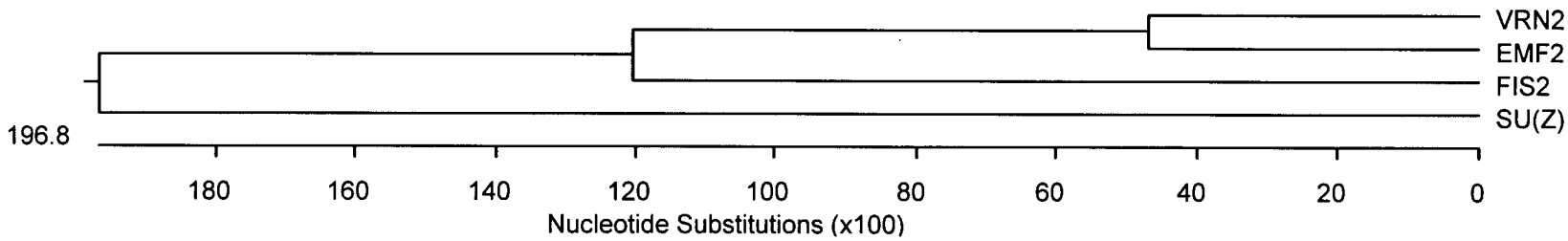


Figure 1.11 FIS2 Phylogenetic tree. Generated using Clustal W (Thompson *et al.* 1994).

1.4.6 A fourth FIS *gene:* MSI1

Recently, Khöler *et al.* (2003) identified a fourth mutant, *msi1*, displaying phenotypes similar to *fie*, *mea*, and *fis2*. *MSI1* encodes a WD-40 protein that is widely conserved in eukaryotes and participates in various complexes involved in chromatin dynamics. Members of this family include p55, a subunit of the protein complex CHROMATIN ASSEMBLY FACTOR (CAF-1) involved in histone loading in *Drosophila* (Tyler *et al.* 1996), the retinoblastoma-associated proteins RbAp46 and RbAp48 in vertebrates (Qian and Lee 1995, Nicolas *et al.* 2000), CAC3p in *Saccharomyces cerevisiae* (Kaufman *et al.* 1997), and LIN-53 in *Caenorhabditis elegans* (Lu and Horvitz, 1998).

Arabidopsis has five *MSI1*-like genes (*MSI1-5*) (Ach *et al.* 1997, Kenzior and Folk 1998, Hennig *et al.* 2003), but only *msi1* has been well characterized. *msi1* mutant gametophytes initiate endosperm development in the absence of fertilization, and plants heterozygous for null *msi1*alleles show a seed abortion ratio of 50% with seeds aborting when the mutant allele is maternally inherited, irrespective of a paternal wild type or mutant *MSI1* allele (Khöler *et al.* 2003). *MSI1* is thought to have an additional role in the adult plant as reducing MSI1 levels by co-suppression causes ectopic expression of floral homeotic genes and strongly interferes with cellular differentiation in seedlings and adult plants (Hennig *et al.* 2003).

1.4.7 Targets of the *FIS* genes

Transcriptional profiling of *mea* and *fie* mutants carried out by Köhler *et al.* (2003) has identified a possible direct target for the FIS products (Köhler *et al.* 2003). One of the genes that showed elevated expression levels in both *mea* and *fie* mutant seed encodes a novel MADS-box transcription factor, subsequently named PHERES1 (PHE1). To confirm the micro-array results, *PHE1* expression was characterised by *in situ* mRNA localisation and PHE::GUS reporter gene studies. These revealed that *PHE1* is not expressed in the wild

type female gametophyte, but is present in the early embryo and endosperm shortly after pollination and is restricted to the CZE later in development. In *fis* mutants unlike wild type, expression was also seen in unfertilized embryo sacs in the central cell 3 days after emasculation, correlating with autonomous endosperm development. After fertilization *PHE1* expression is stronger and more extensive in *mea/fie* mutants that in wild type, consistent with the micro-array results. Köhler *et al.* (2003) also demonstrated that FIE and MEA proteins are associated with the *PHE1* promoter region, suggesting that *PHE1* is a direct target of FIE and MEA. Expression of an antisense *PHE1* transgene, driven by the *MEA* promoter resulted in *PHE1* downregulation and subsequent rescue of *mea* seed. The rescued seed were larger than wild type, remained green longer, and showed reduced tolerance of desiccation (Köhler *et al.* 2003). These results suggest that the lethality of the *fis* mutants is in part due to mis-regulation of *PHE1* in early seed development, and that *PHE1* probably affects endosperm proliferation given the correlation of expression in the central cell with autonomous endosperm development.

1.5 Polycomb-group genes

1.5.1 *Epigenetic control during development*

Building a multicellular organism is a highly complex process, whereby the differentiation and organisation of cells into tissues and organs is not only achieved by following a prescribed genetic program; there is a lot of communication between parts of the organism, and the environment can also influence the developmental process. Plants differ from animals in that organgenesis is continuous and not restricted to the embryonic stage. Also, the body plan is not pre-determined, but can flexibly change according to the interactions of the organism with its environment, and is strongly influenced by external stimuli such as light and nutrients.

Regulation of development in both plants and animals is achieved at many levels to ensure the correct spatial and temporal expression of developmental genes such as the homeotic genes that control developmental patterning and regional identity. The first level of regulation is at transcription, for example, through chromatin structure and remodelling to alter accessibility of *cis*-acting control regions to the transcriptional machinery and associated factors, and the composition and availability of the transcription factors that recruit the transcriptional machinery to the proximal promoter. Chromatin remodelling can involve the non-covalent modification by ATP-dependent remodelling factors that influence DNA/nucleosome contact; or can involve covalent modifications of histone tails (including acetylation, phosphorylation, methylation, and ubiquitination), all of which can affect gene activity by allowing or obstructing the transcription factors or the basal transcriptional machinery access to the *cis*—acting regulatory regions. DNA methylation can also result in chromatin remodelling, as factors that specifically bind ^{m}CpG can then recruit chromatin remodelling factors. Furthermore, it has been shown that transcriptional gene silencing in *Drosophila* requires functional RNAi machinery, originally identified for its requirement in post transcriptional gene silencing (PTGS) (Pal-Bhadra 2002). This suggests further complexity, for example PTGS may be a trigger for transcriptional gene silencing. In plants and animals, epigenetic modifications such as chromatin structure or DNA methylation allows for mitotically and/or meiotically heritable yet reversible changes in gene expression without changes in DNA sequence. This reversibility is important as epigenetic marks can be erased and reset during gametogenesis resulting in a totipotent zygote. Regulation of gene expression also occurs post-transcriptionally, for example, through the processing and stability of mRNA, by translational control, or *via* protein processing and trafficking.

During development, the stable maintenance of expression patterns of homeotic genes is critical for cell fate specification to occur correctly (Bowman *et al.* 1989, Coen *et al.* 1991, Lawrence *et al.* 1994). The regulation of homeotic gene expression has been best studied in

Drosophila where expression patterns are initially set up by the gap and pair-rule proteins. However, these proteins are only transiently expressed, and disappear after about 4 hours following fertilization (Lawrence *et al.* 1994). The continued correct expression of the homeotic genes requires a further two groups of genes; the trithorax-group (trx-G), which act to maintain expression in those cells where the homeotic genes have been turned on, and the Polycomb-group (Pc-G), which act to maintain transcriptional repression of the homeotic genes where they have been initially turned off (Pirrotta 1998). However, this classification may be oversimplified as some Pc-G proteins behave as activators in certain contexts (LaJeunesse 1996). Cell memory is regulated by a balance between these repressors and activators, allowing the transcriptional status to be maintained. In *polycomb* mutant embryos the initial specification of homeotic gene expression is normal, but in later stages of embryogenesis and adult development, ectopic expression of the homeotic genes occurs once gap and pair-rule factor expression diminishes (Orlando *et al.* 1998). Maintaining the repression initiated by the gap and pair-rule factors is thought to be achieved by the formation of a Pc-G complex that modulates chromatin structure to create a repressed state (Pirrotta 1998). It is thought that both Pc-G and TrxG complexes bind to specialised, switchable modular DNA elements called Polycomb response elements (PREs) and core promoters maintaining repression over a wide region (Pirrotta 1998, Lyko and Paro 1999). However, in mammals and plants no PRE-like DNA elements have yet been defined.

In *Drosophila* more than 15 Pc-G genes have been characterised, and encode proteins that are in most cases structurally unrelated. Different Pc-G members have been shown to associate in multimeric protein complexes. Recently, two distinct functional core groups have been distinguished; PRC1 (Polycomb Repressive Complex 1) (Shao *et al.* 1999) a ~3 MDa complex containing Polycomb (Pc), Posterior sex combs (Psc), polyhomeotic (ph), sex combs on midleg and several other proteins, and a functionally separate 600KDa complex,

PRC2 (Polycomb Repressive Complex 2) that contains Extra Sex Combs (ESC) (homologous to *FIE*), Enhancer of zeste (E(z)) protein (homologous to *MEA*), Su(z)12 (homologous to *FIS2*), Pleiohomeotic (PHO), P55, and the histone deacetylase RPD3, (Tie *et al.* 1998 and 2001, Müller *et al.* 2002, Orlando 2003). However, the compositions of the core PRCs differ amongst cell types and extensive changes in compositions occur during the embryonic development of cells (Otte and Kwaks 2003).

1.5.2 *ESC and E(z)*

The WD40 Pc-G gene *ESC* is distinct from other Pc-G genes given that it is expressed and required during early embryogenesis only, unlike the other Pc-G genes, which are required persistently. This suggests a possible involvement for *ESC* in the initiation of silencing, perhaps by recruiting other Pc-G members into repressor complexes on DNA (Simon *et al.* 1995). ESC has been shown to directly interact with the Pc-G SET domain protein E(Z) (Tie *et al.* 1998) and a phosphorylated form of ESC preferentially associates with E(z) *in vivo* (Ng *et al.* 2000). It is thought that the formation of an ESC-E(Z) complex at PREs is an essential prerequisite for the maintenance of silencing. It has been shown that the E(z) SET domain methylates histone H3 at K9 and K27, with a strong preference for the latter (Cao *et al.* 2002, Czermin *et al.* 2002, Kuzmichev *et al.* 2002, Müller *et al.* 2002). It is thought that the E(z)-ESC complex acts in the early embryo in combination with the cell fate determination system, and that the specific H3 K27 methylation marks the target site for binding of the long-term memory PRC1 complex *via* Pc (Orlando 2003). The PRC1 complex is thought to act in a non-catalytic fashion by occluding TrxG proteins. Interestingly, plants do not appear to possess an equivalent long-term memory complex, perhaps reflecting the plasticity of plant development compared to animal development. For example, in plants Pc-G repression has to be released at some point during normal development in the response to certain stimuli that allows progression throughout the life cycle, such as fertilization or flower induction.

1.5.3 *The ESC/E(z) Partnership Shows Striking Evolutionary Conservation*

As well as homologues in *Arabidopsis*, *E(z)* and *ESC* are conserved throughout the plant (Springer *et al.* 2002) and animal kingdoms. The mammalian orthologues of *E(z)* and *ESC* have been identified and termed *Enx1/EZH2* and *EED* respectively (Schumacher *et al.* 1998, Sewalt *et al.*1998); they are implicated in multiple processes including early embryonic patterning, HOX gene regulation, and hematopoiesis (Lessard *et al.* 1999, O'Carroll *et al.* 2001, Schumacher *et al.* 1996). Mouse ESC and E(z) homologues associate with the inactive X chromosome in trophoblast stem cells, suggesting a direct role in X-inactivation (Mak *et al.* 2002, Wang *et al.* 2001).

In fact, vertebrates contain homologues representative of all the Pc-G members so far characterised in *Drosophila.* However, to date the only Pc-G genes found in *C. elegans* with homology to the *Drosophila* genes are the *E(z)* and *ESC* orthologs *MES-2* and *MES-6* (Holdeman *et al.* 1998, Kelly *et al.* 1998, Seydoux *et al.*1999). The *MES* genes (*Maternal Effect Sterile*) were first identified in screens for maternal-effect mutations that result in sterile offspring. The normal nuclear localisation of MES-2 requires wild type *MES-6* function and *vice versa*, supporting the hypothesis that the MES-2 and MES-6 proteins may function as a complex (Holdeman *et al.* 1998, Kelly *et al.* 1998, Seydoux *et al.* 1999). Whereas in both *Drosophila* and mammalian systems the E(z)/ESC complex acts in anterior-posterior patterning in the soma, in nematodes these proteins are primarily involved in germline proliferation and transgene silencing in the germline, and have little effect on the soma.

EED has been shown to interact both *in vitro* and *in vivo* with histone deacetylase (HDAC) proteins (van der Vlag *et al.* 1999). Acetylation of histones is correlated with transcriptional activity, whilst deacetylation is associated with inactive genes. The histone deacetylase

inhibitor trichostatin A relieves transcriptional repression mediated by EED. This supports the idea that deacetylation of histones by HDACs and the recruitment of EED to the HDAC proteins (or recruitment of HDACs to EED) may be among the initial repressive events during embryogenesis that eventually lead to stable and heritable Pc-G mediated repression of target genes.

The ESC/E(z) partnership is therefore widely conserved, appearing in mammalian, *Drosophila*, *C. elegans*, and *Arabidopsis* systems. Although distinct developmental roles can be seen between the species, the partnership between the two proteins as repressors at the level of chromatin appears to be conserved. The ESC and E(Z) proteins are more widely conserved than other members of the *Drosophila* and vertebrate Pc-G (Holdeman *et al.* 1998 and Korf *et al.* 1998) suggesting that these two proteins are of more ancient origin than other Pc-G proteins. Because plant and animals are thought to have evolved multicellularity independently, an ancient Pc-G-dependent mechanism for chromatin modification has been recruited to independent developmental processes in the two kingdoms.

1.6 *Pc-G* genes in *Arabidopsis*

So far three groups of Pc-G proteins with distinct biological functions have been identified in *Arabidopsis*. The first group consists of the *FIS* genes that act during seed development described previously. The second group includes *CURLY LEAF (CLF)* and *EMBRYONIC FLOWER2 (EMF2)*, which act to control aspects of floral induction and floral organogenesis. The third group is required for the vernalisation response, an acceleration of flowering that occurs following long periods of cold treatment.

1.6.1 The *Arabidopsis CLF*, *CLK* and *EMF* genes repress flowering

Arabidopsis flowers contain 4 concentric whorls of organs: sepals, petals, stamens, and carpels (whorls 1-4 respectively). The identity of each whorl is determined by the specific combination of homeotic genes expressed (figure 1.12) (Coen and Meyerowitz 1991, Weigel and Meyerowitz 1994). Most of the floral homeotic genes encode putative transcription factors that share a highly conserved motif called the MADS-box. In the revised ABC model, the combination of class A, B and SEP functions specify petal identity. The combined class B, C and SEP activities specify stamen development; and the presence of the *AGAMOUS* (AG)/SEP complex in whorl 4 is required for the specification of carpel fate. Sepal identity is thought to be specified through the complex of the class A proteins and other yet unidentified factors (Goto *et al.* 2001). In this model A and C functions are mutually antagonistic; in the absence of the C activity it is predicted that class A function expands into whorl 3 and 4, and likewise in the absence of class A function, the C function becomes active in all 4 floral whorls.

A recessive mutation, designated *curly leaf (clf)* (Goodrich *et al.* 1997) was identified during transposon mutagenesis and conferred a phenotype with striking similarity to those of transgenic plants in which the homeotic C-function *AG* gene was constitutively expressed (Mizukami and Ma 1992 and 1997). Ectopic expression of *AG* in the two outer floral whorls results in homeotic transformations of the floral organs in these two whorls. The ectopic expression of AG in *clf* flowers occurs late in floral development, after the onset of AG expression in whorls 3 and 4 (Goodrich *et al.* 1997). This would suggest that the wild type CLF protein is not required for setting up the initial *AG* expression pattern. Interestingly, the CLF protein shows extensive homology to the *Drosophila* Pc-G protein E(Z) (figure 1.9). Thus, CLF seems to perform a similar function to animal Pc-G

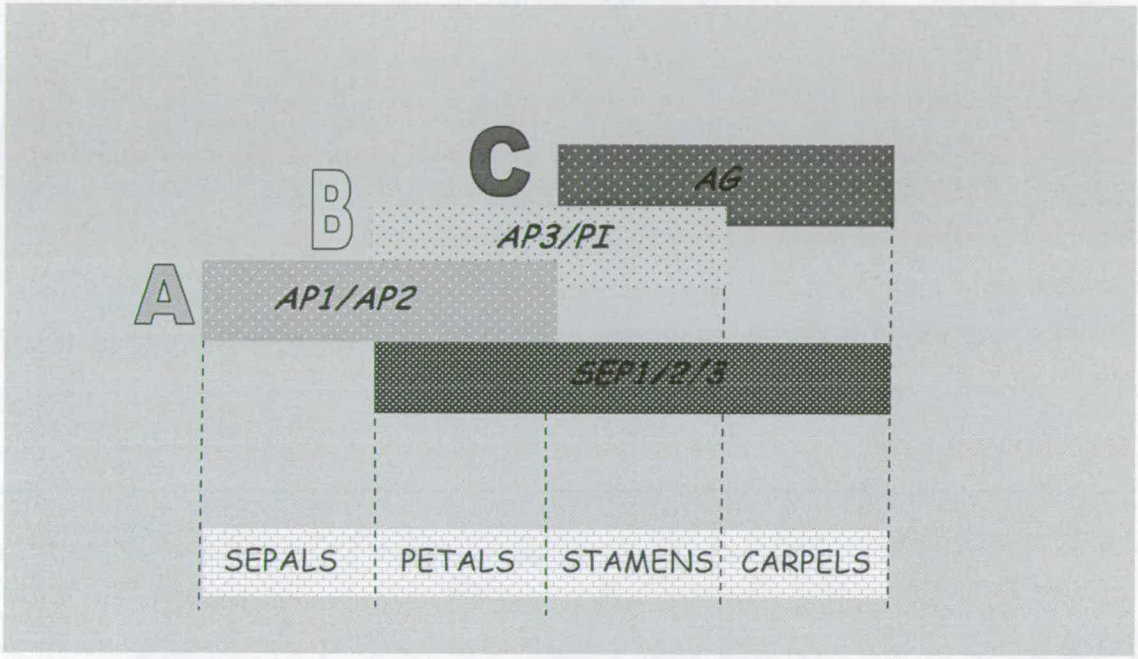


Figure 1.12 The ABC model for specification of floral organ identity. Three overlapping homeotic functions (A, B and C) specify floral organ identity. The combinations of functions active in a whorl specify its identity, e.g. AB specifies petal. A-function genes *AP1* and *AP2* are expressed in whorls 1 and 2, B-function genes *AP3* and *PI* in whorls 2 and 3, and C-function gene *AG* whorls 3 and 4. *SEP1/2/3* function in whorls 2-4. *AP1 (APETALA1), AP2 (APETALA2), AP3, (APETALA3), PI (PISTILLATA), AG (AGAMOUS), SEP1-3 (SEPALLATA1-3),* SE (SEPALS), PE (PETALS), ST (STAMEN), CA (CARPEL). Adapted from Wolpert *et al.* Principles of Development. Current Biology Limited (1998).

proteins, acting to maintain, rather than set up homeotic gene expression patterns in flowers (Goodrich *et al.* 1997).

The CURLY LEAF-LIKE1 (CLK1)/ EZA1 protein shows very high sequence and structural homology to CLF. As well as their expression patterns being very similar, double mutant analysis reveals a synergistic enhancement of the single mutant phenotypes, suggesting that the two genes may be functionally redundant (Y.Chanvivattana, unpublished). The EMBRYONIC FLOWER (EMF) genes EMF1 and EMF2 are required to maintain vegetative development and repress flower development. Loss-of-function mutations in these genes cause *Arabidopsis* to flower directly, bypassing vegetative shoot growth. These mutants show a progressive increase in the expression of flower organ-specific genes (Moon *et al.* 2003). *EMF1* encodes a novel protein that is predicted to be a transcriptional regulator (Aubert *et al.* 2001). EMF2 encodes a zinc finger protein similar to FIS2 and the *Drosophila* protein Su(z)12 (Yoshida *et al.* 2001) (figure 1.10).

1.6.2 The *Arabidopsis* Pc-G gene *VRN2* mediates the vernalisation response

The Pc-G gene *VERNALISATION2* (*VRN2*) encodes a zinc finger transcription factor similar to *FIS2*, *EMF2* and *Su(z)12* (figure 1.10). *VRN2* was isolated in a screen for mutants that do not respond to vernalisation (a period of prolonged exposure to cold temperature that accelerates flowering in temperate biennial plants such as *Arabidopsis*). Although vernalisation occurs at the seedling stage, the transition to flowering happens several weeks later. So the plant needs to remember the stimulus at a later stage in development. Therefore, an epigenetic basis has been proposed for the cellular memory mechanism to explain the vernalisation response. This is thought to involve a Pc-G complex containing VRN2, which targets the MADS-box gene *FLOWERING LOCUS C (FLC)*. In a late flowering *fca* mutant background, cold treatment results in stable repression of *FLC*. Plants

with mutations in *VRN2* lose the ability to maintain the repressed state of *FLC* expression after cold treatment.

Together, these findings suggest that Pc-G function may be conserved in plants as well as in vertebrates and insects. However, as the targets in plants are MADS box transcription factors, not homeobox (as in animals) it seems possible that Pc-G members have acquired independent target genes during plant evolution. There is also a further distinction in that so far, the Pc-G gene *CLF* appears to act specifically on *AG* and *AP3*, *VRN2* on *FLC* and *MEA* on *PHERES1*, while the *Drosophila* counterparts regulate expression of around 100 targets genes (Goodrich *et al.* 1997).

1.7 A role for *FIE* beyond seed development

FIE, unlike the other *FIS* members is a single copy gene that is expressed persistently, raising the possibility that it is required beyond seed development. Consistent with this, Kinoshita *et al.* (2001) rescued *fie* seed viability by using a modified *FIE* promoter (*pFIE*) to drive *FIE* cDNA expression. However, the transgene expression did not recapitulate the wild type pattern of *FIE* RNA accumulation, thought to be because of a requirement for additional regulatory sequences. Transgenic *pFIE::FIE-GFP* plants showed GFP fluorescence in the central cell nucleus before and after fertilization, which disappeared by the eight-nuclei endosperm, and could not be detected at any later stage of plant development. Transgenic *fie* seed bypassed normal *fie* lethality, and the resulting seedlings were smaller than wild type, did not produce normal rosettes, and the seedling shoot apical meristem (SAM) showed premature flowering reminiscent of *emf2* mutants, and inappropriate formation of shoots and flower-like structures along the root and hypocotyl. These plants demonstrated ectopic expression of the floral meristem identity genes *LFY*

(*LEAFY*) and *AP1* and the floral organ identity genes *AG* and *PI* in embryos and seedlings (Kinoshita *et al.* 2001). These results suggest that *FIE* represses the transition to flowering during embryo development and early during vegetative development when the plant normally produces rosette leaves but no infloresences or flowers.

1.8 Summary and objectives

In plants and animals, cell identity is maintained by epigenetic changes that prevent changes in cell type specific transcription programs, like for example the specific combinations of homeotic gene activities that specify regional fate during development. Central to the cell memory machinery are the Polycomb-Group (Pc-G) genes. The *Arabidopsis* Pc-G genes *FIE, FIS2* and *MEA* are required for ovule and seed development. These three genes confer similar mutant phenotypes in both unfertilized embryo sacs and developing seeds; that is, fertilization-independent endosperm proliferation and the abortion of fertilized seeds that carry a maternally inherited mutant *mea, fis2,* or *fie* allele. The common mutant phenotype could be explained by the three *FIS* wild type gene products acting in a common pathway, whereby one gene activates the expression of another to regulate seed development. Alternatively the three FIS proteins may act as part of a multimeric complex similar to that observed with their animal homologues. It has recently been shown that the FIS proteins belong to one such large complex (Köhler *et al.* 2003), which may include proteins possessing specific DNA-binding activity to target the complex to its target gene(s) (*e.g. PHE1*).

In situ hybridisation data (Spillane *et al.* 2000) show that during early seed development following fertilization, *FIE* is expressed in both embryo and endosperm. Aborting *fie* seed show defective endosperm and embryo. During seed development the embryo relies on the

endosperm for nourishment and signals. The embryo arrest could therefore be a secondary result of effects on the endosperm. Occasionally, *mea/mea* or *fis2/fis2* plants are observed, but never *fie/fie*, suggesting either an additional role for the *FIE* product in the embryo, or that *mea* and *fis2* are less severe due to redundancy with related genes. Therefore, although *FIE* is expressed in both zygotic tissues, it is difficult to define its functions in each.

The aims of this project were therefore to test whether MEA, FIE, and FIS2 proteins have the potential to interact by performing yeast two-hybrid assays and to identify novel factors interacting with FIE that could function in DNA targeting by carrying out a yeast two-hybrid screen using FIE as bait. And thirdly, in order to establish the requirement for *FIE* in the embryo and endosperm, a strategy for targeting expression of *FIE* to specific time phases or regions in the seed was employed.

Chapter 2. Materials and Methods

Chapter 2. Materials and Methods

2.1 Plant Materials

2.1.1 *Arabidopsis thaliana* stock lines

clf-2 line

The strong *clf-2* mutant (Landsberg *erecta* (L*er*) ecotype) is derived from a transposon mutagenesis experiment (Long *et al.* 1993).

clf-50 line

The null *clf-50* line confers a severe *clf* phenotype. This allele was identified in a T-DNA mutagenesis experiment in a Wassilewskija (Ws) ecotype, donated by Dr. E Huala.

fie-2 line

In this mutant a G residue is changed to an A residue at the border of the third intron and is consequently thought to disrupt the splicing reaction (Ohad *et al.* 1999) in ecotype Lansberg *erecta* (L*er*).

Enhancer trap lines

Enhancer trap lines 9194, 9185, and k22 (ecotype C24) expressing *GAL4-VP16* and *UAS::mGFP5* were generated by Jim Haseloff and colleagues.

Salk lines

Salk lines 017005 and 061267 (Columbia) carrying a T-DNA insertion at the *RAP-1* locus were generated by the Salk Institute Genome Analysis laboratory (SIGnAL) and obtained from the *Arabidopsis* Biological Resource Centre (ABRC).

2.1.2 Plant growth conditions and pollination

i. Seed Sterilisation

Vapour method

Three ml of hydrochloric acid was added to 100ml of concentrated sodium hypochlorite solution to generate chlorine gas in a sealed chamber. The seed were then treated for 8-16 hours.

Liquid method

Seeds were washed in 20ml concentrated sodium hypochlorite solution, 20ml dH_2O, one drop of tween for two minutes, followed by 70% ethanol for five minutes. The seeds were then rinsed twice in sterile dH_2O.

After sowing, imbibed seeds were stratified by keeping at 4^0C for 3-5 days to promote and synchronise germination.

ii. Growing *Arabidopsis* on soil

Plants were grown on a 3:1 mixture of Levingston's compost: grit in a greenhouse, under long days, or under short or long day photoperiod at 23°C in growth rooms.

iii. Growing *Arabidopsis* on tissue culture plates

Sterilised seed were sown on petri-dishes containing 1% agar, 4.604g/l MS salt (Sigma), 10g/l glucose, 100g mg/l inositol, and 0.5g/l MES pH 5.7. Marker selection was made by adding appropriate antibiotics (50 µg/ml kanamycin or Hygromycin-B). Seedlings were transferred onto soil after 2-3 weeks.

iv. Cross-pollination

Female parents were emasculated by removing immature anthers from floral buds. Pollination was carried out the next day on open flowers by smearing pollen from the anthers of the male parent onto the mature stigma of the emasculated female flowers.

2.2 DNA Analysis

2.2.1 Isolation of DNA

i. Isolation of Plant DNA

Alkali lysis method for PCR analysis

A small piece ($3mm^2$) of rosette or young cauline leaf was put in an eppendorf tube containing 50µl 0.25M sodium hydroxide. The tube was boiled for 30 seconds after which 50µl of 0.25M hydrochloric acid and 25 µl of 0.5M Tris-HCl pH8.0 was added and the tube and boiled for a further 2 minutes. The tissue was then used directly in PCR reactions.

Small-Scale Genomic DNA Extraction for PCR Analysis

0.5 ml CTAB extraction buffer containing 100 mM Tris-HCl pH 8, 1.4 M NaCl, 20 mM EDTA, 2% (w/v) CTAB and 0.2% 2-mercaptoethanol (Rogers and Bendich, 1985) was heated to 60°C. 1-2 leaves (rosette or young cauline), which had previously been ground in liquid nitrogen, were added and the mixture incubated at 60°C for 20-30 minutes. Proteins were extracted by addition of 5 ml of chloroform. The organic phase was separated by centrifugation (13,000 rpm for 15 min). The top (aqueous) layer was removed and the DNA precipitated by addition of 2/3 volume of isopropanol. DNA was collected by centrifugation (13,000 rpm, 15 min), washed in 70% (v/v) ethanol, left to air dry and dissolved in TE buffer (10 mM Tris—HCl pH 8 and 1 mM EDTA pH 8).

Large-Scale Genomic Extraction for Southern Analysis

5-10 leaves were ground into a fine powder with a pre-cooled pestle and mortar in liquid nitrogen. The tissue was then transferred into a chilled eppendorf tube containing 500µl extraction buffer (50mM EDTA, 0.1M NaCl, 0.1M Tris-HCl pH8, 1% SDS). An equal volume of phenol/chloroform was added and the sample vortexed, left on the bench for 5 minutes, vortexed again and centrifuged at 13,000 rpm for 5 minutes. The aqueous phase was transferred to a clean tube and the phenol/chloroform extraction step was repeated. DNA was precipitated by adding 50µl 3M sodium acetate pH5.5 and 350µl isopropanol. The DNA was pelleted by centrifuging at 13, 000 rpm for 5 minutes, then rinsed with 70% ethanol, air-dried and resuspended in 50µl TE solution. One third of the DNA extracted was used for Southern analysis.

ii. Isolation of Bacterial Plasmid DNA

E. coli plasmid DNA extraction

Single colonies were inoculated in independent 5ml LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl) containing the appropriate antibiotic. Cultures were incubated overnight at 37^{0}C with constant agitation. The next day bacterial cells were harvested from 1.5ml of culture in an eppendorf tube by centrifugation at 13,000 rpm in a microcentrifuge for 1 minute. Cells were then resuspended in a 100µl GTE solution (50mM glucose, 25mM Tris-HCl pH 8, 10mM EDTA, 100µg/ml Rnase A). To lyse the cells 200µl 0.2M NaOH/1%SDS solution was added and the tubes were inverted several times to mix the solutions. This was followed by adding 150µl ice-cold 3M KOAc pH 4.8 and repeated inversions to mix. After centrifugation for 5 minutes at 13, 000rpm in a microcentrifuge, the clear supernatant was carefully transferred into a fresh eppendorf tube containing 1ml of ethanol. After mixing by inverting the tubes several times, the DNA was pelleted by centrifugation for 5 minutes at

13,000rpm in a microcentrifuge, washed in 70% ethanol, air-dried and resuspended in 50µl TE.

Agrobacterium plasmid DNA Extraction

Single colonies were inoculated into 5 ml of LB containing appropriate antibiotics, and then incubated shaking at 28°C until culture reached stationary stage (overnight to 3 days). Bacterial cells were harvested from 1.5ml of culture in an eppendorf tube by centrifugation at 13,000 rpm in a microcentrifuge for 1 minute then resuspend in 100µl lysis buffer (50mM Glucose, 25mM Tris-Cl pH 5.8,10mM EDTA, 4mgml^{-1} Lysozyme) and incubated at room temperature for thirty minutes. Then 200µl 0.2M NaOH/1% SDS solution was added and mixed by gentle shaking, and incubated at room temperature for 10-30 min. Then 150µl of cold 3M NaOAc, pH 4.8 was added and mixed by shaking and the tube was incubated on ice for 5 minutes. The tubes were then centrifuged for 5 min in a microcentrifuge at 13,000rpm. The supernatant was transferred to a fresh tube containing 0.7 volumes isopropanol and mixed. The DNA was then pelleted by centrifugation at 13,000 rpm for five minutes, washed with 70% ethanol and resuspended in 10µlTE.

2.2.3 Restriction Analysis of DNA

DNA was digested in 20µl reactions containing [1X] appropriate restriction buffer and 3.5units/µg substrate DNA of the corresponding enzyme as supplied by the manufacturers. The samples were incubated for 60-180 minutes at 37^{0}C or temperature recommended by manufacturers. Genomic DNA for Southern analysis was digested in [1X] restriction buffer containing 1mM spermidine for at least six hours at 37^{0}C.

2.2.4 DNA Gel Electrophoresis

Electrophoresis was performed in agarose gels submerged in TBE (2.5mM EDTA, 90mM boric acid, and 90mM Tris-HCl pH8.0). Gels were cast 4-8mm thick to determine the size and presence of DNA and to separate DNA fragments for purification. Gels used were 0.65-2.5% agarose (Bioline) in TBE. Loading buffer (0.0025% bromophenol blue, 0.0025% xylene cyanol and 0.25% ficoll type 400) was added to the DNA samples to be loaded. Electrophoresis was carried out at 8 volts/cm for 1-3 hours. For southern blot analysis digested genomic DNA was run in 0.7% agarose gels at 1-2 volts/cm for 20-72 hours. Gels were stained in ethidium bromide (10μg/ml) for 20-30 minutes and the DNA was visualised under UV light.

2.2.5 Purification of DNA from Agarose Gels

The desired DNA fragments were identified using a low energy UV transilluminator before being excised from the agarose gel using a clean scalpel. DNA was then extracted and purified using the Qiagen QIAquick extraction kit or the QIex extraction kit as described in the corresponding manufacturer's manuals.

2.2.6 Southern Hybridisation

A Digoxygenin labelled probe was used for salk_061267 (*rap-1-2*) DNA, and a 32pCTP-labelled radioactive probe for salk_017005 (*rap-1-1*), spanning the full-length *RAP-1* cDNA.

Radiolabelling of DNA probe

The DNA fragment (500ng) was added to 60ng random hexamer primer, boiled for 2 minutes, spun briefly and put on ice. A labelling mix was added containing 5mM Tris-HCl pH6.9, 1mM $MgSO_4$, 10μM dithioreitol, 6 μM dATP, 6 μM dTTP, 6 μM dGTP, 50 μCiα32-P dCTP and 5 units of DNA polymerase I (Klenow fragment, Roche Scientific) in a 20 μl

reaction. The reaction was incubated for at least 1 hour at room temperature. To separate the probe from unincorporated nucleotides, the reaction mix was applied onto a 1ml Sephadex G50 column followed by the addition of 180 µl TE pH8.0 and 20µl 10X NT dye (1% Dextran Blue, 0.1% Orange G). The blue fraction containing the longer probe fragments was collected and heated at 100^{0}C prior to hybridisation to denature the probe.

i. Southern Hybridisation with Digoxygenin-Labelled probes

Digoxygenin probe labelling

A 50µl PCR reaction was set up containing 0.5 µl of primer1 [0.5 µg/µl] (m13 forward) and primer2 (m13 reverse), 5 µl Boehringer PCR digoxygenin labelling mix (2mM dATP, 2 mM dCTP, 2mM dGTP, 1.3 mM dTTP, 0.7 mM alkali-labile DIG-11-dUTP, pH 7.0), 1X PCR buffer (Boehringer), 2ng/µl template DNA (pGemTeasy *RAP-1* $^{-299 \text{ to } +2029 \text{ (relative to ATG)}}$), 1 unit of *Taq* polymerise, and made up to 50 µl with dH_2O.

Hybridisation

Ethidum bromide-stained gels were soaked in 0.25M HCL with gentle agitation for 15 minutes to depurinate large DNA fragments and aid DNA transfer during blotting. This was followed by a 30-minute wash in denaturation solution (0.5M NaOH, 1.5M NaCl) and two 15-minute washes in neutralization solution (0.5M Tris-HCL pH7.2, 1.5M NaCl, 1mM EDTA). DNA was transferred onto Hybond-N (Amersham) Nylon membrane on a wick platform with 20 x SSC (1M NaCl, 100mM sodium citrate) as the transfer buffer. After 15-20 hours of transfer, the filters were washed briefly in 2 x SSC solution, air-dried then UV cross-linked (40J/cm^2) to fix the DNA onto the membrane. Before hybridisation, the filters were incubated at 65^{0}C in 70ml of hybridisation buffer (5 x SSC, 0.02% SDS, 0.1% N-lauryl sarcosine and 1% (w/v) dry milk) for two hours. The solution was then replaced with 30ml of fresh hybridisation solution containing the denatured digoxgenin labelled probe.

Hybridisation was carried out overnight at 65^{0}C with gentle shaking. The filter was then washed twice at room temperature in 2 x SSC, 0.1% SDS, followed by two washes in 0.5 x SSC, 0.1% SDS for 15 minutes at 68^{0}C. The filter was then incubated in blocking solution (100mM maleic acid, 150mM NaCl pH 7.5 1% w/v blocking powder) for one hour at room temperature. The blocking solution was then replaced with 20ml fresh blocking solution containing 2µl antiDIG AP Fab antibody, and incubated at room temperature for 30 minutes. The filter was then washed twice for 15 minutes at room temperature in 100mM maleic acid, 150mM NaCl pH 7.5 0.3% tween. The filter was then equilibrated in detection buffer (100mM Tris-HCl pH 9.5, 100mM NaCl) for 3 minutes at room temperature. The filter was then soaked in chemiluminescent substrate solution (20µl CSPD in 2ml detection buffer), and sealed in a bag and incubated at 370C for 10 minutes before exposure to X-ray film for 20 minutes to 16 hours.

2.2.7 Preparation of DNA Samples for Sequencing

DNA sample (40ng/1kb) was mixed with dH2O to a volume of 4.4 µl in a PCR tube. To each tube was added 1.6 µl of 1 µM sequencing primer and 4 µl of BigDye™ v3.0 sequencing mix (ABI PRISM®). PCR was then performed using the following cycles:

96°C for 30 seconds
96°C for 30 seconds
50°C for 15 seconds
60°C for 4 minutes
} 25 cycles

The extension products were then transferred to an eppendorf tube and DNA was precipitated by adding 1 µl 3M sodium acetate pH5.5 and 25 µl ethanol. The tube was incubated on ice for 10 minutes then centrifuged for 15 minutes at full speed. The DNA was washed in 70% ethanol before being air-dried. The samples were analysed on an ABI377 sequencer by Jill Lovell (ICAPB).

2.2.8 Polymerase Chain Reaction (PCR)

i. PCR reaction

The PTC-200 DNA engine (MJ Research) was used to carry out the PCR reactions. A typical 50µl reaction contained approximately 50ng of template DNA, 1X PCR buffer (1mM Tris-HCl, 1.5mM $MgCl_2$, 50mM KCl, pH8.3), 0.4 µl each of forward and reverse primer (100mM), 200 µM of each dNTP, and 0.2 units of *Taq* DNA polymerase. Reactions were preheated to 94°C before the polymerase was added. PCR was then typically performed using the following cycles:

94°C denaturation for 2 minutes
94°C denaturation for 2 minutes
55°C-65°C annealing for 30 seconds
72°C extension for 1 minute per kb
72°C extension for 10 minutes
} 30-40 cycles

The primers used in the PCR reactions are shown in table 2.1.

ii. RT-PCR reaction

600ng oligo-dT primer (Stratagene) was used in a 10µl reaction containing 50mM Tris-HCL pH 8.3, 75mM KCl, 10mM DTT, 3mM MgCl2, 1mM each dNTP, 1 unit Rnase inhibitor (Stratagene) and 50 units MMLV reverse transcriptase (Stratagene) and incubated at 37°C for 30 minutes. 1/10 of the sample was then used as a template in a standard PCR reaction.

Primer ID	Sequence (5' to 3')
CM21	aat atc gaa tgt cga cga taa cct tag gga a
CM21RC	ttc cct aag gtt atc gtc gac att cga tat t
CM22	gcc ttg tta acg gcg gcc gct gta cca cag tg
CM22RC	cac tgt ggt aca gcg gcc gcc gtt aac aag gc
MEA1	cga gtg gtt aat gtc gac gga aaa cca tga g
MEA2	ctc atg gtt ttc cgt cga cat taa cca ctc g
MEA3	ttg atc tga gga gcg gcc gca att caa gca

MEA4	tgc ttg aat tgc ggc cgc tcc tca gat caa
FIEBGL-A	gac aga gag aga tct ttc gaa tat cg
FIEBGL-B	cga tat tcg aaa gat tct tct ctg tc
FIENCO-A	acg tga tta cca tgg agc ggt ctg agt
FIENCO-B	act cag acc gct cca tgg taa tca cgt
FIE5'SAL	aga gag aga gat gtc gac tat cga atg tcg
FIE5'SAL RC	cga cat tcg ata gtc gac atc tct ctc tct
MEAEcoRIRC	ctt ctc cat taa gaa ttc gcc tcc ccc ccc
MEAEcoRI	ggg ggg gga ggc gaa ttc tta atg gag aag
CPS3	gca aaa ccc acg tca ttt cag tgc ctc
CPS3RC	gag gca ctg aaa tga cgt ggg ttt tgc
CPS4	ctc ttg aag aag acg tcc cat tat ttc ttg
CPS4RC	caa gaa ata atg gga cgt ctt ctt caa gag
CPS5	gat tta tat gga cgt ctg ggc agg act atg
CPS5RC	cat agt cct gcc cag acg tcc ata taa atc
CPS8	tca ttt cag tgc gtc gac tga tta taa ttc
CPS8RC	gaa tta taa tca gtc gac gca ctg aaa tga
CPS9	caa atc act ttt gtc gac aat gca aaa ccc
CPS9RC	ggg ttt tgc att gtc gac aaa agt gat ttg
CPS10	aca cca tca atc gtc gac gtt aaa cca gcc
CPS10RC	ggc tgg ttt aac gtc gac gat tga tgg tgt
46830FC	ttg atg atc gaa gct ctc ctt g
46830RC	tga ggg acg acg ttt tga at
46760FC	tga acg gca caa cat cat ca
46760RC	tcg cag ctg atc taa cca ca
fis2xhoFC	atg atg ata aag tct cga gga cac caa gag ctc ac
RAP-1/A	ggt cac cgg ttt atg gaa tg
RAP-1/B	gcc gga tct tgg ttt cac ta
LBb1	gcg tgg acc gct tgc tgc aac t
RAP1-3	cga tac aga ctc aaa cat aga gca aaa ga
RAP1-5	gta cga tct cac tca ctt ccg gtc tta at
RAP1-zig5	ggt cac cgg ttt atg gaa tg
RAP1-zig3	aaa att act gaa gcc act aaa acg a
salk border	gct gtt gcc cgt ctc act ggt g
m13f	gta aaa cga cgg cca gt
m13r	gga aac agc tat gac cat g
FIS2A	taa gat tca act cga ggg atg tga gta ac

FIS2B	ctc gag atc tcc gcg gcc gct aac cac tg
EZA	cgt gtc aag tcg acg tac atc aaa atc
EZB	tgg gca cct ttt ggc ggc cgc tgg gaa tg
RAP1int1	tgc tct att ttg ttt tca atg tgg
RAP1int5	tgg ttc ccc tta aca ttg tct c

Table 2.1 PCR Primer sequences

2.2.9 *In Vitro* Mutagenesis (IVM)

Base changes were incorporated into (miniprep) plasmid DNA using the QuikChange® site-directed mutagenesis kit (Stratagene). The QuikChange® kit can be used to make point mutations, switch amino-acids, and delete or insert single or multiple amino-acids. Two oligonucleotide primers are designed to contain the desired mutation, each complementary to the opposite strands of the plasmid containing the target site. The oligonucleotides are then extended during temperature cycling by (the high fidelity) *Pfu Turbo* DNA polymerase, to generate a mutated plasmid containing staggered nicks. The product is then treated with *Dpn I* to digest the parental (dam methylated) DNA template. The nicked vector is then transformed into XL1-blue *E.coli* cells, amplified and recovered.

2.2.10 Ligation of DNA fragments

Ligations were carried out in a 10µl total volume, in 1 X ligation buffer (Roche Scientific) containing 1 unit of T4 DNA ligase (Roche Scientific). Vector and insert DNA was added in an approximate molar ratio of 1:1. Reactions were then incubated at 14^0C overnight.

2.2.11 Generation of Blunt-Ended DNA Fragments by 3′ Recessed End Fill-In

1μg DNA was dissolved in 50mM Tris pH 7.2, 20mM MgSO4, 200μM DTT, 40μM dNTPs with 1 unit of Klenow (NEB), in a total volume of 100μl. The reaction was incubated for 15 minutes at 25°C, and then stopped by heating at 75°C for 20 minutes.

2.2.12 Transformation of *E.coli* cells with DNA Plasmid

i. Transformation using electro-competent *E.coli* cells

Cell preparation

A one-litre culture of *E.coli* DH5α strain was grown in LB medium from a 10ml overnight culture at 18^{0}C. When an OD_{600} 0f 0.5-0.7 was reached, cells were incubated on an ice water bath for 30 minutes, then centrifuged at 4500rpm in a chilled Sorval GSA rotor for 15 minutes at 4^{0}C. The cell pellet was washed in one litre ice cold 10% glycerol, followed by further washes in 500ml, 100ml, 20ml, and 10ml ice cold 10% glycerol and finally resuspended in 2-3ml ice cold 10% glycerol. Aliquots of 100μl were transferred into eppendorf tubes and frozen in liquid nitrogen before being stored at –80^{0}C.

Electro-transformation

Frozen competent cells were thawed on ice; 40μl of cells were mixed with 1μl ligation reaction in a pre-chilled transformation cuvette. Cells were shocked with an 18 KV electric pulse in a Gene Pulser machine (Biorad) at 200 Ω and 25μF and immediately resuspended in 1ml SOC media (2% bactotryptone, 0.5% bacto yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$ and 20mM glucose). The cells were then incubated at 37^{0}C with constant shaking for one hour, after which the cells were pelleted by centrifugation at 10,000rpm in a microcentrifuge for 30 seconds, resuspended in 100μl LB medium and spread out on ampicillin (100μg/ml) or kanamycin (50 μg/ml) LB agar plates (plates for

blue/white screening also contained X-gal (40 μg/ml) and IPTG (200 μg/ml). Plates were incubated overnight at 37^{0}C.

2.2.13 Subcloning

Yeast-two hybrid constructs

pBI881/pBI880 FIE: primers CM21, CM21RC, CM22, and CM22RC were used to introduce in frame 5' *SalI* and 3' *NotI* restriction sites by Quickchange® site-directed-mutagenesis into the *FIE* cDNA containing plasmid pCMesc3. *FIE* cDNA was then subcloned in-frame into the yeast-2-hybrid vectors pBi880 and pBi88I (figure 3.2).

pEG202/pJG4-5 FIE: primers were used to introduce in-frame 5' *EcoRI* by Quickchange® site-directed-mutagenesis. The FIE coding sequence was then subcloned into pEG202 and pJG4-5 as an *EcoRI/XhoI* fragment.

pBI881/pBI880 MEA: Primers MEA1, MEA2, MEA3, MEA4 were used in Quickchange® site-directed-mutagenesis to introduce the in-frame 5'*SalI* and 3'*NotI* restriction sites into the plasmid pMEA0. *MEA* cDNA was then subcloned in-frame into the yeast-2-hybrid vectors pBi880 and pBi88I.

pEG202/pJG4-5 MEA: primers MEAEcoRIRC and MEAEcoRI were used to introduce an in-frame 5' *EcoRI* site by Quickchange® site-directed-mutagenesis. The *MEA* coding sequence was then subcloned into pEG202 and pJG4-5 as a *EcoRI/XhoI* fragment.

pBI880ESC/pBI881ESC: primers were used to introduce in-frame 5' *Bgl*II and 3'*NotI* restriction sites by Quickchange® site-directed-mutagenesis, using the template pe2sf

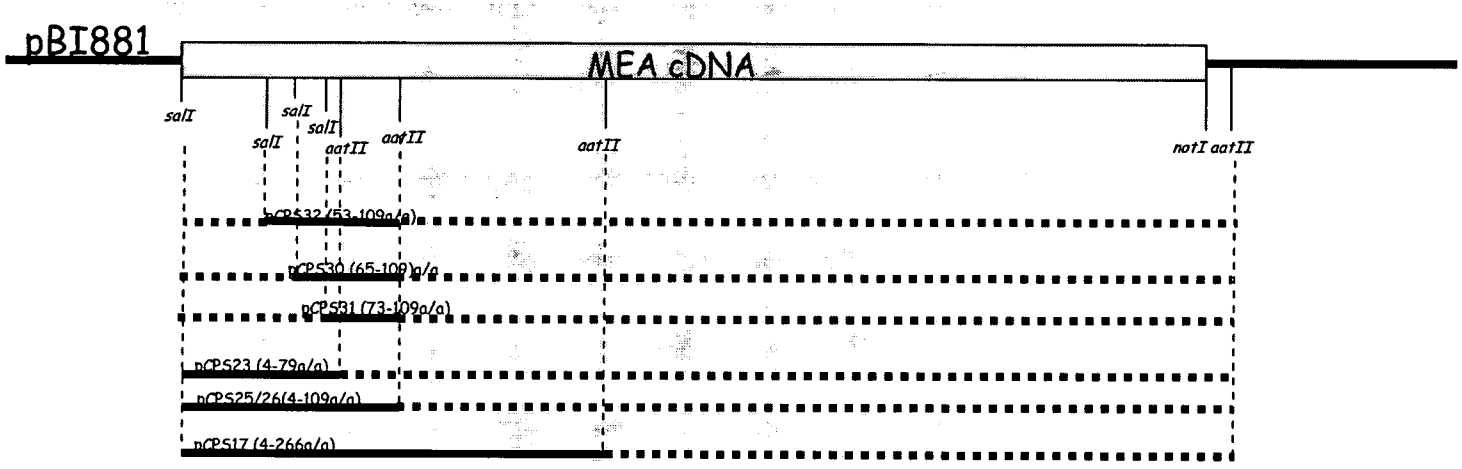
(Simon *et al.* 1995). The *ESC* coding sequence was then subcloned in-frame into the plasmids pBI880 and pBI881.

pBI880E(z)/pBI881E(z): EZA and EZB primers were used to introduce in-frame 5' *Sal*I and 3'*NotI* restriction sites by Quickchange® site-directed-mutagenesis, using the template pe32-55.26 (Jones and Gelbart, 1993). The *E(z)* coding sequence was then subcloned in-frame into the plasmids pBI880 and pBI881.

pBI881MEA deletions: To produce prey constructs that expressed truncated MEA proteins, the full length MEA prey construct pBI881MEA was modified by Quickchange® site-directed-mutagenesis to introduce *Sal*I (amino-terminal truncations) or *Aat*II (carboxy-terminal truncations) sites within the MEA coding sequence (figure 2.1). The modified plasmids were then cut with *Sal*I or *Aat*II to excise 5' or 3' cDNA fragments, and recircularised.

pBI880FIS2/pBI881FIS2: The *FIS2* coding sequence was amplified by RT-PCR. Total RNA was extracted from young siliques (L*er* ecotype) using the plant RNeasy kit (Qiagen) according to the manufacturer's recommendations. 500ng total RNA was reverse transcribed, and used as a template for PCR amplification using the mutagenic primers FIS2A and FIS2B for 40 cycles using a 60^{0}C annealing temperature and the high fidelity *PfuTurbo* polymerase. The product was digested with *Xho*I and *Not*I and ligated to *Sal*I/*Not*I digested pBI880 and pBI881.

pBI881FIS2VEFS: The mutagenic primers fis2xhoFC and FIS2B were used to introduce the 5' in-frame *Xho*I and 3' *Not*I restriction sites by RT-PCR using silique cDNA (see above) as template and the high fidelity *PfuTurbo* polymerase. The purified blunt-ended PCR product was treated with 5 units of *Taq* DNA Polymerase in 1 X PCR Buffer with $MgCl_2$



Plasmid ID	Primer pair ID	Site introduced	Region encoded
pCPS30	CPS8, CPS8RC	*salI*	*65-109*
pCPS31	CPS9, CPS9RC	*salI*	*73-109*
pCPS32	CPS10, CPS10RC	*salI*	*53-109*
pCPS23	CPS4, CPS4RC	*aatII*	*4-79*
pCPS25	CPS5, CPS5RC	*aatII*	*4-109*
pCPS17	-	-	4-266

Figure 2.1 pBI881MEAdeletions. To produce prey constructs that expressed truncated MEA proteins, the full length MEA prey construct pBI881MEA was modified by Quickchange® site-directed-mutagenesis to introduce *Sal*I (amino-terminal truncations) or *Aat*II (carboxy-terminal truncations) sites within the MEA coding sequence. The modified plasmids were then cut with *Sal*I or *Aat*II to excise 5' or 3' cDNA fragments, and recircularised.

containing dATP to a final concentration of 0.2mM, in a final reaction volume of 10µl. The reaction was incubated at 70°C for 15–30 minutes before being ligated into pGEM®-T Easy Vector and subcloned into *SalI*/*Not*I digested pBI881 as a *Xho*I/*Not*I fragment.

pTFTFIE: pCMESC3 was cut with *Xho*I, before being partially cut with *EcoR*I. The fragments were isolated by gel purification and the 1307bp band was ligated to the *EcoR*I/*Sal*I digested vector for yeast 3-hybrid analysis (pTFT). pTFT (a generous gift of Marcos Egea-Cortines, Max-Planck-Institute for Plant Breeding Research) has the ADH1 promoter and nuclear localization signal from pGAD424 (Clontech) as used in the two-hybrid expression plasmids (Fields and Song, 1989). The pTFT vector complements the *ade2* auxotrophy from *Saccharomyces cerevisiae,* and is based on pASZ12, a 2µ version of pASZ11 (Stotz and Linder, 1990). Selection can be made in media lacking adenine.

pTFTMEA: pMEA-0 was cut with *Sal*I and *Xho*I. The *MEA* coding sequence was then ligated into the *Sal*I cut pTFT in the sense orientation.

pBI881RAP-1 deletion series: pBI881RAP-1 was recovered from a yeast two-hybrid screen using FIE as bait, and was found to encode the *RAP-1* N-terminal region (amino acids 234-624). Native internal restriction sites in the *RAP-1* cDNA were used to generate deletion plasmids (figure 2.2). Δ1: pBI881RAP-1 was digested with *EcoR*I/*Not*I, blunt ended, and the linear plasmid was then gel purified before being recircularised. Δ2: pBI881RAP-1 was digested with *Aat*II; the linear plasmid was then gel purified before being recircularised. Δ3: pBI881RAP-1 was digested with *Not*I. The reaction was then incubated at 70°C for 20 minutes to inactivate the enzyme, before being digested with *Hpa*I and blunt-ended. The linear plasmid was then gel purified before being recircularised, by ligation in dilute conditions that favoured intramolecular ligation.

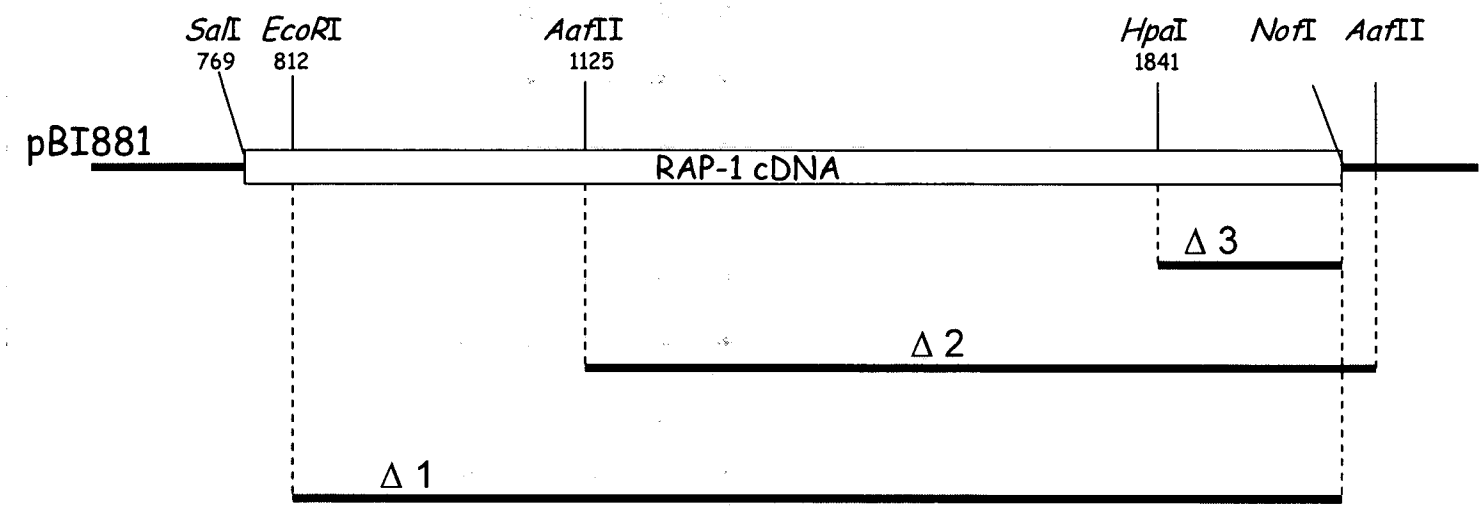


Figure 2.2 pBI881RAP-1deletion series. pBI881RAP-1 was recovered from a yeast two-hybrid screen using FIE as bait, and was found to encode from amino acid 769. Native internal restriction sites in the *RAP-1* cDNA were used to generate deletion plasmids. Δ1: pBI881RAP-1 was digested with *EcoRI*/*Not*I, blunt ended, the linear plasmid was then gel purified before being recircularised. Δ2: pBI881RAP-1 was digested with *Aat*II; the linear plasmid was then gel purified before being recircularised. Δ3: pBI881RAP-1 was digested with *Not*I. The reaction was then incubated at 70°C for 20 minutes to inactivate the enzyme, before being digested with *Hpa*I and blunt-ended. The linear plasmid was then gel purified before being recircularised.

YFP/CFP Fusion Constructs

pMON-CFP-FIE/ pMON-YFP-FIE: The FIE cDNA clone pCMESC3 was mutagenised using the Quickchange® site-directed-mutagenesis kit (Statagene) and the primer pairs FIENCO-A and FIENCO-B, and FIEBGL-A and FIEBGL-B, to introduce a *Bgl*II site upstream of the ATG start codon and a *Nco*I site immediately upstream of the TAG stop codon. The mutagenised *FIE* coding sequence was cut with *Bgl*II and *Nco*I and ligated to the *Bgl*II/*Nco*I digested pMON999-YFP and pMON999-CFP cowpea expression vectors. These vectors contain a multiple cloning site between the double cauliflower mosaic virus (CaMV) 35S promoter and a terminator sequence of the nopaline synthase gene (Tnos) to allow expression of the fusion proteins in plants (van Bokhoven *et al.* 1993).

Binary Construct

pBIB-UAS hyg^R: Using the Stratagene QuikChange® Site-Directed Mutagenesis Kit a suitable restriction site (*SalI*) was introduced into pCMESC3 (containing the full-length *FIE* cDNA) to allow subcloning into the modified binary vector pBIB-Hyg^R (Becker 1990). This was downstream of the previously introduced GAL4 UAS element (figure 2.3). The primer pair FIE5'SAL and FIE5'SAL RC was used to introduce a *SalI* site upstream of the *FIE* translational start site in the plasmid pCMESC3. The *FIE* cDNA was then isolated by digestion with the restriction enzymes *Sal1* and *Kpn1* and ligated into the polylinker of the modified pBIB-HYG vector. The plasmid was then amplified in E-*coli* and subjected to restriction analysis to verify the presence of the FIE cDNA before being introduced into *Agrobacterium* by transformation. After *Agrobacterium* transformation the plasmid was re-isolated and re-introduced into *E-coli* to confirm that no rearrangements had occurred.

Templates for *in situ* probes

pBsSK-Rap-1trunc (spanning +733 to +1011 relative to ATG). The *RAP-1* insert from pBI881RAP-1 was subcloned into pBluescript SK- (Stratagene) as a *SalI*/*Not*I fragment.

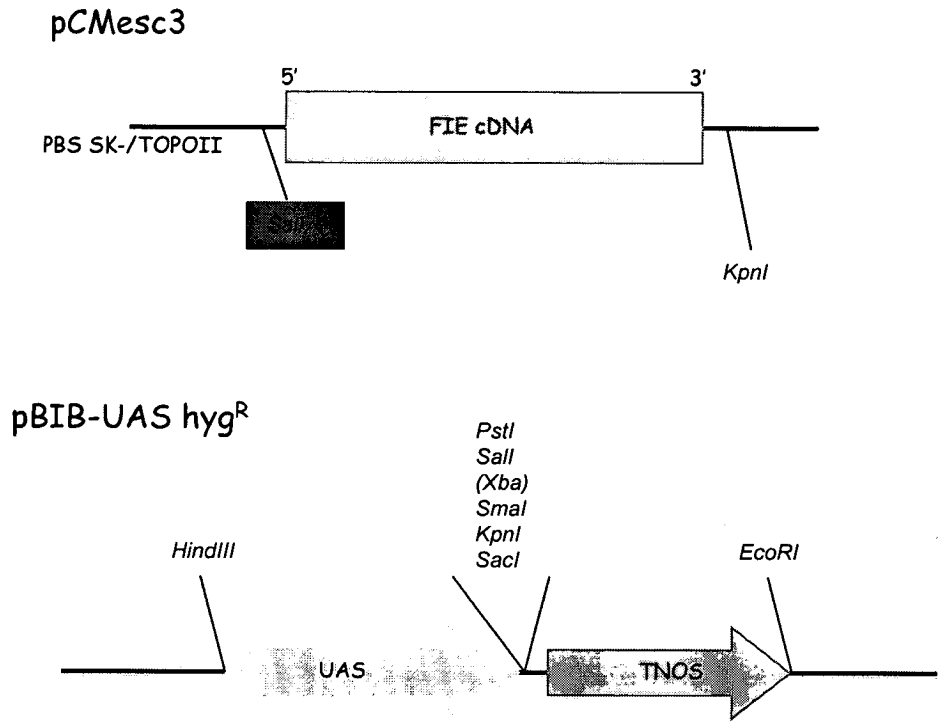


Figure 2.3 pBIB-UAS-FIE Subcloning of *FIE* cDNA into pBIB-UAS binary vector. *SalI* (red) was introduced to pCMESC3 containing the full-length *FIE* cDNA to allow subcloning of the *SalI-KpnI FIE* fragment into the GAL4 UAS containing modified binary vector pBIB-HYG. The UAS element was previously introduced as a *HindIII-SalI* fragment by Gwyneth Ingram (ICMB). TNOS = nopaline synthase terminator sequence.

The C terminal region was removed by restriction with *BamHI* followed by re-circularisation of the plasmid.

pBsSK+Rap-1trunc (spanning +733 to +1011 relative to ATG). The *RAP-1* insert of pBsSK-Rap-1trunc was subcloned directionally into pBluescript SK+ (Stratagene) as a *SalI/SacII* fragment.

2.3 Yeast 2-Hybrid Analysis

i. Yeast transformation

10ml YPD media (1% w/v yeast extract, 2% w/v bacto peptone, 2% w/v glucose) was inoculated with yeast strain YPB2 or AH109 (Clontech) or EGY48 (Clontech) and incubated with agitation overnight at 30^0C. For details of yeast strains see figure 2.4. Cells were counted the next day using a hemocytometer, and added to 100ml fresh YPD media such that a cell concentration of 2 x 10^7 cells/ml was obtained and incubated further. When cells reached a density of 7 x 10^6 the cells were pelleted by spinning in a Sorvall centrifuge for 5 minutes at 7000rpm, washed in 25ml dH_20 and repelleted. Cells were then washed in 1ml 100mM lithium acetate, before being resuspended in fresh 100mM lithium acetate to a final cell density of 2 x 10^9 cells/ml and transferred to eppendorf tubes in 50µl aliquots. To transform the yeast cells each 50µl aliquot was centrifuged at 13,000 rpm for 15 seconds in a microcentrifuge and the lithium acetate removed. 240µl 50% PEG, 36µl 1M lithium acetate, 25µl 2mg/ml single-stranded salmon sperm DNA and 1µg of plasmid DNA was added to the cells. The tubes were then vortexed thoroughly for 1 minute, incubated for 30 minutes at 30^0C, 25 minutes at 42^0C and centrifuged for 15 seconds at 8000 rpm in a microcentrifuge. Cells were resuspended in 200µl dH_20 then plated onto freshly made YMM plates (2% w/v

Strain	System	Genotype[a]	Reporter(s)	Transformation Markers	References
EGY48[b]	LexA	MATa, ura3, his3 , trp1, LexAop (x6)-LEU2	LEU2	his3, trp1, ura3	Estojak *et al.*, 1995
AH109	GAL4	MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, trp1, leu2 gal4D, gal80D, LYS2 : : GAL1 UAS-GAL1 TATA-HIS3, MEL1 GAL2 UAS-GAL2TATA-ADE2, URA3::MEL1 UAS-MEL1TATA-lacZ	HIS3, ADE2, lacZ MEL1	trp1, leu2	James *et al.*, 1996; Holtz, Unpublished
YPB2	GAL4	*MAT*a *ura*3-52 *his*3-200 *ade*2-101 *lys*2-801 *trp*1-901 *leu*2-3, 112 *can*^{R} *gal*4-542 *gal*80-538 *LYS*2::$GAL1_{UAS}$-$LEU2_{TATA}$-*HIS*3 *URA*3::($GAL_{17mersx3}$)-$CyC1_{TATA}$-*lacZ*)	HIS3, lacZ	trp1, leu2, his3	Bartel 1993

Figure 2.4 Yeast Host Strain Information· [a]The trp1, his3, gal4, and gal80 mutations are all deletions; leu2–3, 112 is a double mutation. [b]Also contains the pSH18-34 *lacZ* reporter plasmid.

glucose, 0.67% yeast nitrogen base without amino acids, 2% w/v agar (Difco), amino-acids (see table 2.2)), and incubated at 30^0C for 3 days.

Amino-acids (Sigma)	Concentration (µg/ml)
L-arginine – HCL	20
L-aspartic acid	100
L-glutamic acid	50
Adenine sulphate	*40*
L-histidine –HCL	*20*
L-isoleucine	30
L-leucine	*160*
L-lycine- HCL	30
L-methionine	20
L-phenylalanine	50
L-serine	380
L-theonine	200
L-tryptophan	*40*
L-tyrosine	30
Uracil	20
L-valine	150
Proline	20
Glycine	20
Cysteine	20
Cystine	20
Alanine	20
Glutameine	20
Asparagine	20

Table 2. 2 Yeast media amino-acid concentration. Selective plates were made by omitting the appropriate amino-acids in italics.

ii. Yeast Colony Filter Assay for β-galactosidase Activity

Independent yeast transformants were picked using a sterile toothpick and streaked onto Nylon-backed nitrocellulose filters (Amersham) placed on top of a fresh YMM-L-W plate and incubated at 30°C for 3 days. Filters were submerged for 10 seconds in liquid nitrogen before being saturated with 2ml Z-buffer (60mM Na_2HPO_4, 40Mm NaH_2PO_4, 10mM KCL, 1mM $MgSO_4$, 1mg/ml X-Gal and beta-mercapthoehanol) and incubated at 37°C for between 1hr and overnight.

2.4 *In situ* Hybridisation

The expression pattern of *RAP-1* was analysed by *in situ* hybridisation with labelled RNA probe. *In vitro* transcribed RNA probe was non-radioactively labelled with digoxygenin-11-UTP and hybridised to longitudinal sections of *Arabidopsis* inflorescence tissue. Probe detection was carried out by an immunochemical assay with an anti-digoxygenin antibody conjugated to an alkaline phosphatase. The anti-digoxygenin antibody binds to digoxygenin and the phosphatase enzyme catalyses the colour reaction with Western Blue substrate (Promega) such that the signal can be detected as a blue/purple precipitate.

2.4.1 Tissue fixation and embedding

Inflorescence issues were harvested and fixed in 4% (v/w) paraformaldehyde. Paraformaldehyde was dissolved in 1X phosphate buffered saline (PBS, 130mM NaCl, 7mM Na_2HPO_4, 3mM NaH_2PO_4) that had been warmed to 60°C and the pH adjusted to 11 with NaOH pellets. The solution was then slowly cooled down to 4°C and its pH readjusted to 7 using H_2SO_4. In order to improve the penetration of the fixative into the tissues, 0.1% (w/v) Triton-X-100 and 0.1% (w/v) Tween 20 were added. The plant materials were immersed in the fixative and a vacuum was applied for 15 minutes in order to aid the penetration of the

fixative. The fixative was then replaced with fresh solution and fixation proceeded overnight at 4^{0}C with gentle shaking.

After fixation, the tissue was washed twice in 1X PBS with gentle shaking to remove the paraformaldehyde. Subsequently the tissue was dehydrated through an ethanol series (Table 2.3). All the steps involving paraformaldehyde were carried out in a fume hood. The tissue was then cleared with Histoclear (CellPath). Wax was introduced at the second 100% Histoclear treatment by adding ¼ volume of paraplast chips (BDH- Merke) and leaving overnight in a 42^{0}C heating block. After overnight incubation, a further ¼ volume of melted chips was added and the wax/Histoclear mixture was replaced with fresh molten wax after 2 hours treatment at 60^{0}C. The tissue was incubated in wax overnight at 60^{0}C. The molten wax was replaced twice in the morning and evening of the following day. This was repeated for the next two days. The details of the treatments are listed in table 2.3. Finally the tissue was placed in plastic moulds containing molten wax. The position of the tissue was carefully adjusted for subsequent sectioning. The plastic mould was cooled rapidly to solidify the wax and later stored at 4^{0}C until sectioning.

Treatment	Duration	Temperature
30% ethanol	1 hour	4^{0}C
40% ethanol	1 hour	4^{0}C
50% ethanol	1 hour	4^{0}C
60% ethanol	1 hour	4^{0}C
70% ethanol	1 hour	4^{0}C
85% ethanol	1 hour	4^{0}C
95% ethanol	Overnight	4^{0}C
100% ethanol	30 minutes	room temperature

100% ethanol	30 minutes	room temperature
100% ethanol	1 hour	room temperature
100% ethanol	1 hour	room temperature
75% ethanol, 25% Histoclear	1 hour	room temperature
50% ethanol, 50% Histoclear	30 minutes	room temperature
25% ethanol, 75% Histoclear	30 minutes	room temperature
100% Histoclear	1 hour	room temperature
100% Histoclear	1 hour	room temperature

Table 2.3 Treatment for embedding tissue

2.4.2 Sectioning of embedded tissues

The wax block containing the fixed tissue was cut using a razor blade into a trapezoid, mounted onto a small metal disc and at 4^0C for at least 15 minutes. The disc was secured onto a microtome (Reichet-Jung) and its position adjusted such that the face of the trapezoid was aligned next to and parallel to the blade. Sections ware cut into ribbons of 8μM thick and floated on sterile water on a Probe On Plus coated slide (Fisher Scientific). The slide was placed in a 42^0C hot plate for about 15 minutes until the ribbon flattened. Water was removed by suction with a sterile Pasteur pipette and blotting with a lens cleansing tissue (Whatman). Finally the slides were left at 42^0C overnight to allow the ribbons to adhere onto the slide surface. The slides were then stored at 4^0C until use.

2.4.3 *In vitro* transcription of RNA probes

The desired *RAP-1* DNA fragment to be transcribed was subcloned into the polylinker of pBluescript KS^- (Stratagene) next to the T7 polymerase promoter (for details see section 2.blah). Plasmid DNA was then amplified in *E.coli* before being extracted and purified. Using the T7 primer the insert was PCR amplified then purified from agarose using the

QIAGEN QIAquick PCR purification kit, dried in a vacuum concentrator (Savant) and 0.5-1μg was resuspended in 18μl RNase-free water. The diluted DNA was added to a labelling reaction mix containing 2.5μl 10 X RNA polymerise buffer (Roche Scientific), 1μl RNase inhibitor (20 units/ μl), 2.5μl 5mM ATP, GTP and CTP, 2.5 μl 1mM DIG-UTP and 1μl RNA polymerise (20 units/μl). The reaction was incubated at 37^{0}C for 2 hours and then stopped by adding 75μl RNase-free water, 1 μl 100mg/ml tRNA and 3μl of DNase I (10 units/μl). After a 15 minute incubation at 37^{0}C to allow DNase I to digest template DNA, the RNA was precipitated using 1 volume of 3.5M sodium acetate and 2 volumes of ice-cold ethanol. The RNA pellet was washed in DEPC-treated 70% ethanol, air-dried and resuspended in RNase-free water. The successful incorporation of the DIG-UTP during the labelling reaction was estimated by spotting 1μl of the newly labelled probe and 1/100 dilution of the DIG-labelled control RNA (Roche Scientific) onto a nylon filter. The RNA was UV crosslinked at 0.4 J/cm^2. The filter was then wetted with buffer 1 (100mM Tris-HCl pH7.5, 150mM NaCl), incubated with gentle shaking in buffer 2 (1 X Blocking agent (Roche Scientific) dissolved in buffer 1) for 30 minutes, washed briefly with buffer 1, and then incubated in buffer 1 with 1μl of anti-DIG antibody. The filter was then washed twice in buffer 1 for 15 minutes to remove unbound antibody and then briefly washed in buffer 5 (100mM Tris-HCl pH9.5, 100mM NaCl, 50mM $MgCl_2$) to raise the pH. Finally, the filter was immersed in Western Blue substrate (Promega) and shaken gently for 15 to 30 minutes until blue/purple spots became visible. The relative intensities of the RNA probe spot and the control spots gave an estimate of the success of the transcription reaction.

2.4.5 Hybridisation

i. Section pre-treatment

Slides containing sectioned tissues were loaded onto a pre-baked stainless steel rack. Slides were treated with a series of solutions as shown in table 2.4. All the steps were carried out in

RNase-free plastic tubs and 300ml of solution was used in each step (600ml for the acetic anhydride treatment). The Histoclear and the first ethanol steps were performed in a glass container. Since acetic anhydride is unstable in water, fresh acetic anhydride was added to the solution with vigorous mixing using a magnetic stirrer with the slide rack elevated above the stirrer on an empty rack. Apart from the proteinase K digestion which was carried out at 37^{0}C, all the treatments were performed at room temperature.

Treatment	Duration (minutes)
Histoclear	10
Histoclear	10
100% ethanol	2
100% ethanol	2
95% ethanol	2
90% ethanol	2
80% ethanol	2
60% ethanol	2
30% ethanol	2
DEPC-treated water	2
2 X SSC	20
Proteinase K (1µg/ml in 100mM Tris-HCl pH8.0, 50mM EDTA)	30
Glycine (2mg/ml in PBS)	2
PBS	2
PBS	2
Acetic anhydride (0.5% v/v in 0.1M triethanolamine, pH8.0)	10
PBS	2
PBS	2
Rehydrated through ethanol series as above (30% ethanol - 100% ethanol)	0.5 each

Table 2.4 Section pre-treatment

ii. Hybridisation

Pre-treated slides were dried completely of ethanol on clean paper towels. The digoxygenin-labelled RNA probe was thawed on ice and 2µl of the probe was diluted in 19µl of DEPC-treated water. The diluted probe was added to 20µl of formamide, denatured for 2 minutes at 80^{0}C and then added to the hybridisation mix (100µl 10 X *in situ* salt (3M NaCl, 100mM Tris-HCl pH8, 100mM NaPOP4 pH6.8, 50mM EDTA), 400µl deionised formamide, 200µl 50% dextran sulphate, 20µl 50 X Denhardts solution (1% polyvinylpyrrolidone, 1% Ficoll, 1% BSA fraction V), 10µl tRNA (100mg/ml) and 70µl DEPC-treated water). Two slides were sandwiched together and the probe mix was allowed to coat the slides by capillary action. The slide pairs were placed on top of glass rods in a plastic tray saturated with 50% formamide. The tray was then sealed and incubated at 55^{0}C overnight.

2.4.6 Post-hybridisation washes

After overnight hybridisation, the slide pairs were separated by rinsing in 0.2 X SSC solution then placed in a clean rack. The rack was then subjected to a series of washes with gentle agitation as shown in table 2.5.

Treatment	Temperature	Duration
0.2 X SSC	Room temperature	1 hour
0.2 X SSC	Room temperature	1 hour
NTE	37^{0}C	5 minutes
NTE	37^{0}C	5 minutes
RNase (20µg/ml in NTE)	37^{0}C	30 minutes
NTE	37^{0}C	5 minutes
NTE	37^{0}C	5 minutes
0.2 X SSC	55^{0}C	1 hour
PBS	Room temperature	45 minutes

Table 2.5 Slide washes

2.5.7 Detection of bound digoxygenin labelled probe

The slides were washed in 1% Roche Block (in buffer 1). After 45 minutes the blocking solution was replaced by buffer 3 (1% BSA in 100mM Tris-HCl pH7.5, 150mM NaCl. 0.3% Triton-X-100) and the slides washed for another 45 minutes. Following the washes digoxygenin-AP antibody (Roche Scientific) was diluted 1:1250 in buffer 3. The slides were sandwiched together again and the antibody solution was drawn up by capillary action. The solution was drained off and fresh solution was reapplied. The slide pairs were then placed on glass rods in a plastic tray and incubated at room temperature for 2 hours. The antibody was then drained off the slides and the slides were separated, returned to the rack and washed 4 times for 15 minutes each in buffer 1 with gentle shaking. This was followed by a 10-minute wash with buffer 5. Slides were then re-sandwiched and capillary action was used to draw Western Blue substrate up. The first application was drained on a clean paper towel and the solution was reapplied. The slide pairs were elevated on glass rods in a sealed container and placed in total darkness for 2-3 days. Subsequently the slides were separated, placed back in the rack and immersed in TE buffer pH8.0 for 15 minutes to stop the colour reaction.

2.5 Microscopy

DIC microscopy

Developing seeds and ovules were mounted in a solution containing chloral hydrate: dH_2O: glycerol in an 8g: 2ml: 1ml ratio. After 1 to 18 hours the slides were examined using a Nikon (Tokyo, Japan) microscope using differential interference contrast (DIC) optics.

Confocal microscopy

Developing seeds and ovules were mounted in a 0.4-0.5% agarose solution and observations were made using an inverted laser scanning confocal microscope (model LSM510; Zeiss,

Jena, Germany) with a Q 63 plan objective. GFP fluorescence was monitored with a 505±550 nm band-pass filter (488 nm excitation line of a krypton±Argon laser).

Fluorescent microscopy

Seedlings were mounted in dH_2O and observations were made using an Axioskop 2 microscope (Carl Zeiss, Jena, Germany). Chloroplast autofluorescence was eliminated with a red emission filter.

Chapter 3. Protein-protein interactions mediating the *fis* phenotype

Chapter 3. Protein-protein interactions mediating the *fis* phenotype

3.1 Introduction: A FIS complex

The *FIS* genes regulate cell proliferation both in the female gametophyte and during early seed development in *Arabidopsis* (Sørensen *et al.* 2001). The *FIS* genes: *FIE, MEA*, and *FIS2* encode Polycomb-group (Pc-G) proteins, transcriptional regulators that form multimeric complexes and are thought to modify chromatin structure to confer a mitotically stable repressed state (Grossniklaus *et al.* 1998a, Grossniklaus *et al.* 1998b, Kiyosue *et al.* 1999, Luo *et al.* 1999, Ohad *et al.* 1999, Tie *et al.* 1998, Jones *et al.* 1998, Shao *et al.* 1999, van der Vlag *et al.* 1999). *FIE, MEA*, and *FIS2* encode homologues of the *Drosophila* E(Z), ESC, and Su(z)12 proteins, respectively. The three proteins are widely conserved in the animal kingdom, and the animal homologues associate in a multimeric complex that is biochemically distinct from the PRC1 Pc-G complex (Sewalt *et al.* 1998, Jones *et al.* 1998, Shao *et al.* 1999, van der Vlag *et al.* 1999, Korf, *et al.* 1998, Holdeman *et al.* 1998, Van Lohuizen *et al.* 1998, Ng, *et al.* 2000, Sewalt *et al.* 1998 Birve *et al.* 2001). To test whether the similarities in *fis* mutant phenotypes also reflect interactions between their protein products, FIS protein interactions using the yeast two-hybrid assay and their protein localization *in planta* were characterized.

3.2 FIE and MEA interact in yeast two-hybrid

The common mutant phenotype (section 1.4 and 1.5) of *mea*, *fie*, and *fis2* could be explained by the three wild type gene products acting in a common pathway, whereby one gene activates the expression of another to regulate seed development. Alternatively the three FIS proteins may act as part of a multimeric complex similar to that observed with their animal

homologues (figure 3.1). To test whether MEA, FIE, and FIS2 proteins have the potential to interact, yeast two-hybrid assays were performed.

The yeast two-hybrid system (Fields and Song 1989) takes advantage of the modular properties of eukaryotic transcription factors such as GAL4 or LEXA (Ptashne 1996). These proteins contain two discrete functional domains, a DNA binding domain (DB) and a transcriptional activation domain (TA). The DB domain is sufficient to bind its specific target sequence, for example, the upstream activating sequence (UAS) in the case of GAL4, but alone it cannot activate transcription of target genes. By contrast TA domains lack DNA binding activity, but are sufficient to confer transcriptional activation, most likely by stabilising interactions between the basal transcription machinery and the target gene promoter. Interaction between two proteins of interest (X and Y) can be detected by expressing DB-X (bait) and TA-Y (prey) fusion proteins in recombinant *Saccharomyces cerevisiae*. If X and Y interact, the TA domain is recruited to the DB fusion, reconstituting GAL4 or LEXA activity, allowing transcription of the reporter genes bound by the DB fusions. To monitor interaction, two reporter constructs are commonly used, comprising of the *Lac-z* coding sequence encoding β-galactosidase (β-gal) and the *HISTIDINE3* (*HIS3*) gene encoding an enzyme required for histidine synthesis. These genes are placed under the control of promoters containing GAL4 or LEXA target sequences in a histidine auxotrophic yeast strain.

The full-length FIE and MEA proteins were expressed as fusions using a GAL4 based yeast 2-hybrid system designed by Dr W. Crosby (NRC Plant Biotech Inst, Saskatoon, Canada) (Kohalmi *et al.* 1998). In this system, the vector pBI880 is used to engineer 'bait' fusions to the GAL4 DB, and pBI881 for 'prey' fusions to GAL4-TA (figure 3.2). The resulting constructs were introduced into the yeast strain YPB2 (Bartel 1993) to express FIE

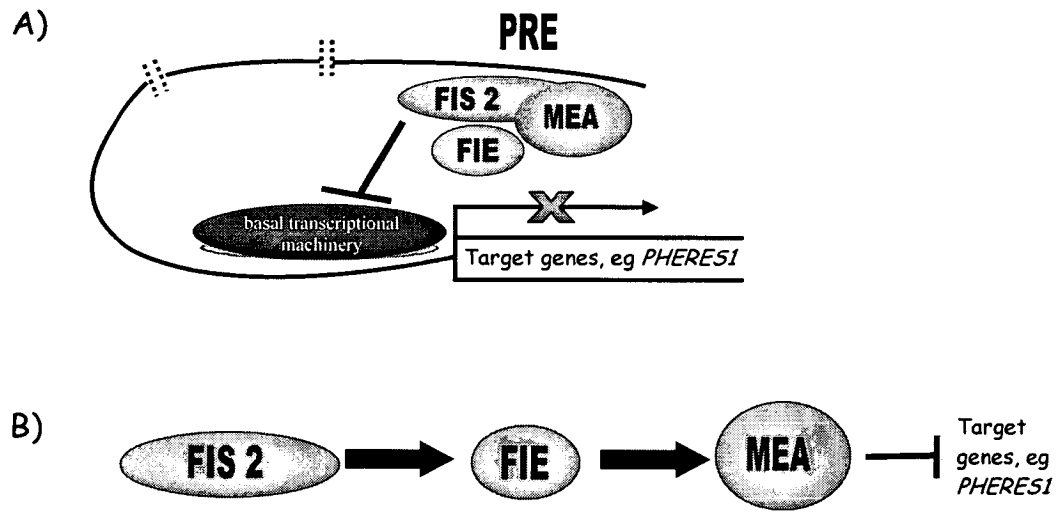
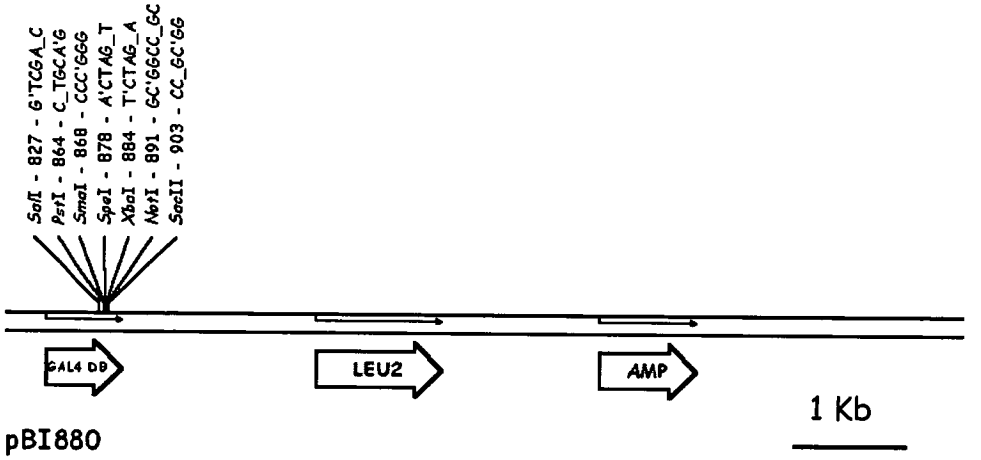


Figure 3.1 Alternative Models for FIS Protein Action. A) Shows the three FIS proteins acting as a complex to repress target genes. B) Shows the three FIS proteins acting in a common pathway to repress target genes. PRE, Polycomb Response Element.

A)



B)

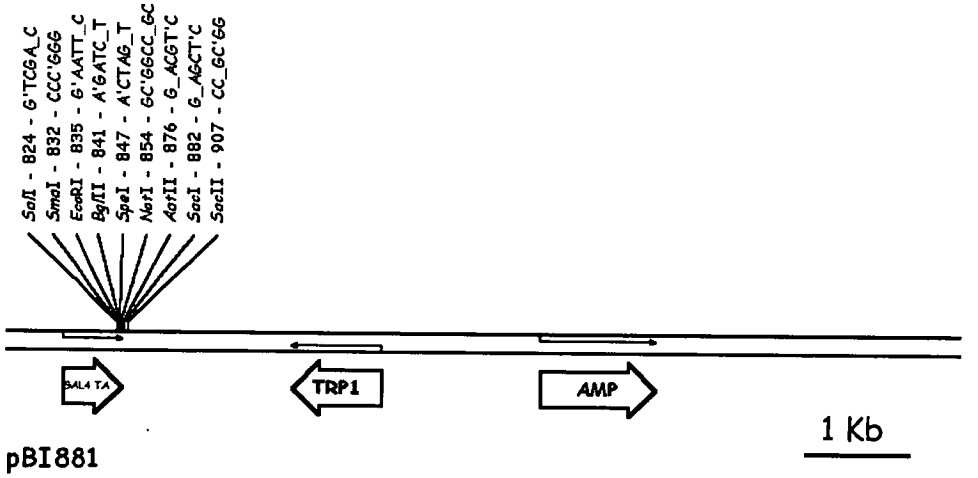


Figure 3.2 Vectors used for Yeast two-Hybrid Assay. A) pBI880 vector for 'bait' fusions. B) pBI881 vector for 'prey' fusions. cDNAs were introduced as in-frame fusions using engineered *Sal*I and *Not*I restriction sites. The vectors can be propagated in *E.Coli* and selected by ampicillin resistance. The *LEU2* and *TRP1* genes allow for selection in yeast strains that carry *leu2* or *trp1* mutations.

and MEA as bait fusions to the yeast GAL4-DB domain and as prey fusions to the GAL4-TA domain. The yeast strain YPB2 is auxotrophic for leucine and tryptophan, which allows for selection of the bait and prey plasmids respectively. Transformed yeast strains that carry the FIE bait construct alone are unable to grow in the absence of histidine and do not express β-galactosidase, indicating that the FIE protein does not provide transcriptional activation when fused to the GAL4 DB (figure 3.3). In addition, yeast containing FIE–DB and GAL4–TA failed to activate either reporter, indicating that GAL4-TA alone does not interact with FIE-DB (figure 3.3). However, yeast transformants carrying both *DB–FIE* and *TA–MEA* express both reporters suggesting that FIE and MEA interact (figure.3.3). The interaction is not detected when the reciprocal fusions (FIE–TA and MEA-DB) were co-expressed (data not shown). One possible explanation is that one or both fusions were poorly expressed; alternatively the interaction domain of one of the proteins may have been masked or improperly folded when fused to the GAL4 domain. The interaction was also tested in an alternative LEXA based two-hybrid system that uses the yeast host strain EGY48 (Golemis *et al.* 1996), in which *LEU2* and *lacZ* reporter genes are transcriptionally dependent on reconstituted LEXA activity. In the EGY48 host, expression of the prey construct is galactose inducible, so colonies are grown on glucose or galactose containing media to compare reporter gene activation in the presence of bait alone or bait and prey, respectively. In this system MEA–DB constructs on their own are able to weakly activate reporter gene expression, but show increased β-galactosidase activity in the presence of FIE–TA prey (figure 3.4) also consistent with a FIE–MEA interaction.

A further confirmation that FIE and MEA proteins can interact was obtained from *in vitro* pull-down experiments performed by C. Spillane as part of a collaboration. Here, FIE protein was expressed *in vitro* using a coupled transcription/translation system. The MEA protein was expressed as a fusion to Glutathione-S-transferase (GST) in *E.coli*, and purified from bacterial extracts. When the two proteins were mixed, and GST-MEA bound to

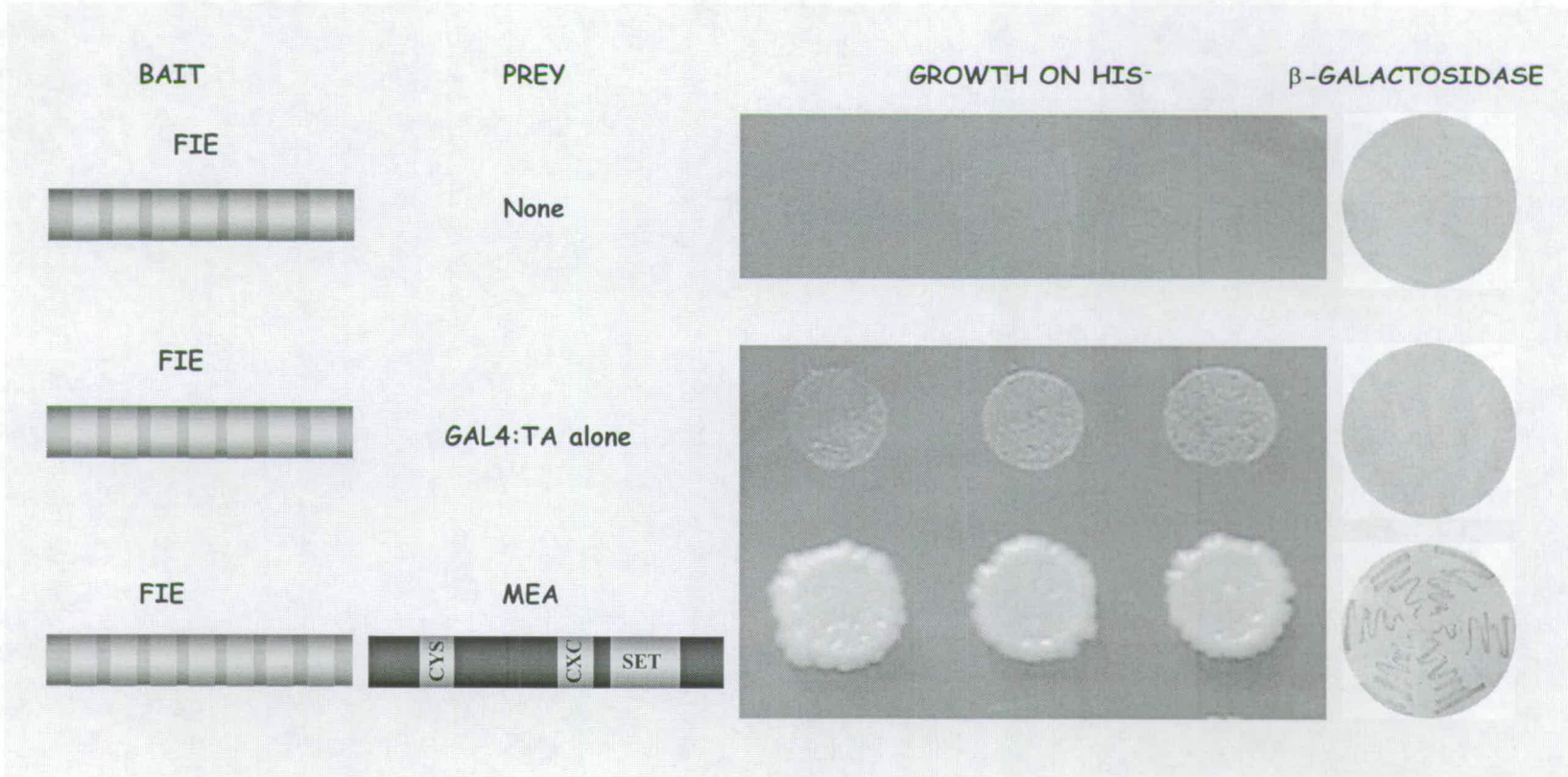


Figure 3.3 FIE Interacts with MEA in yeast. *HIS3* and *lacZ* reporter genes are transcriptionally dependent on reconstituted GAL-4 activity. Three independent transformants are shown for each growth assay on His- media. For ß-galactosidase activity assays six independent transformants are shown on filter paper after lysis and incubation with X-GAL substrate.

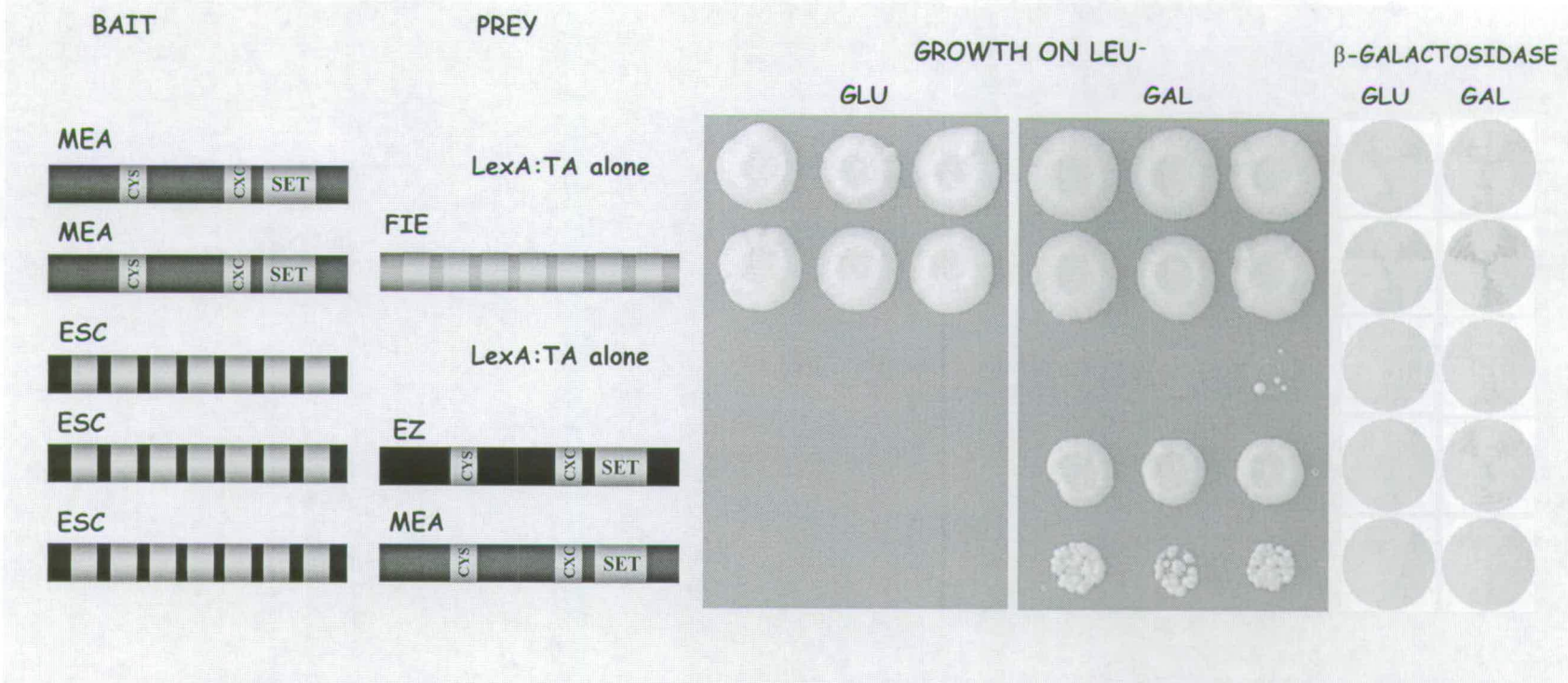


Figure 3.4 *Drosophila* ESC shows a weak interaction with MEA in yeast. LEU2 and lacZ reporter genes are transcriptionally dependent on reconstituted LEXA activity. Expression of the prey construct is galactose inducible, so colonies were grown on glucose or galactose containing media to compare reporter gene activation in the presence of bait alone or bait and prey, respectively. Three independent transformants are shown for each growth assay on Leu- media. For ß-galactosidase activity assays three independent transformants are shown on filter paper after lysis and incubation with X-GAL substrate. MEA-DB alone activated the LEU2 reporter gene but gave weaker activation of the ß-galactosidase reporter. FIE-DB alone strongly activated both reporter genes and therefore could not be used in this assay.

glutathione beads, FIE was also recovered (Spillane *et al.* 2000). These results further support a direct interaction of FIE and MEA.

3.3 Mapping the region of MEA that interacts with FIE

To test whether a discrete region of MEA is sufficient for interaction with FIE, a series of carboxy-terminal truncations of the MEA prey were tested with FIE bait in yeast two-hybrid assays. The results indicated that amino-terminal residues 4–109 of MEA are sufficient for interaction with FIE (figure 3.5). Thus, prey constructs with more extensive deletions of MEA (residues 4–79) are unable to interact with FIE (figure 3.5). In addition, a series of amino-terminal deletions (53-109, 65-109 and 73-109) all tested negative for interaction with FIE. Together these results suggest that residues 4–109 of MEA are sufficient for interaction with FIE.

An analogous deletion series of FIE was not attempted as it has been shown that even small deletions within the WD region of its *Drosophila* homologue ESC drastically affects its structural integrity and greatly diminishes binding to E(z) (Tie *et al.* 1998). However, to test whether mutations that eliminate FIE activity *in planta* also affect its interaction with MEA, the *fie-6* mutation was introduced into the *FIE–GAL4–DB* bait. *fie-6* confers a strong mutant phenotype equivalent to that of alleles encoding extensive truncations or deletions (Ohad *et al.* 1999). It encodes FIE W364STOP which truncates the FIE protein by six residues. This disrupts the highly conserved WD residues at the carboxy-terminal end of WD repeat motif seven, which are predicted to stabilize the β-propeller structure (for example Tie *et al.* 1998, Jones *et al.* 1998 and Ng *et al.* 1998). Introduction of the W364STOP mutation into the FIE bait prevents its interaction with MEA (figure 3.5), consistent with the interaction of FIE and MEA being necessary for FIE activity *in planta*.

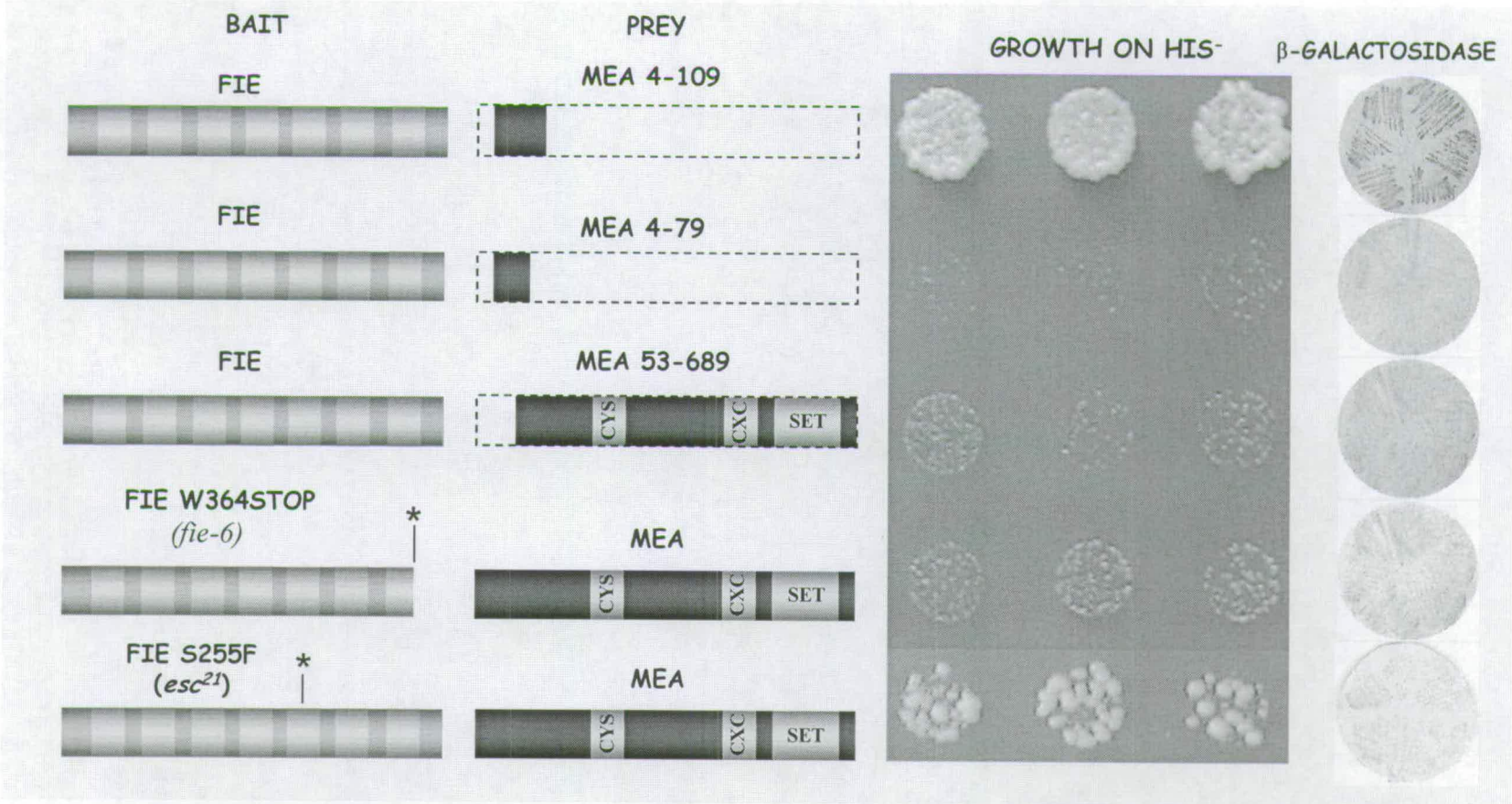


Figure 3.5 FIE Interacts with the N-terminus of MEA in yeast. *HIS* and *lacZ* reporter genes are transcriptionally dependent on reconstituted GAL-4 activity. Three independent transformants are shown for each growth assay on His- media. For ß-galactosidase activity assays six independent transformants are shown on filter paper after lysis and incubation with X-GAL substrate. Only the most informative deletions are shown here. A series of C-terminal deletions (4-109, 4-266) all tested positive, whereas 4-79 tested negative. All the N-terminal deletions (53-109, 65 -109 and 73 -109) tested negative.

The *Drosophila* mutation esc^{21} (S301F in WD5 of ESC) has a drastic effect on E(Z) binding (Tie *et al.* 1998). The serine residue affected is one of the three highly conserved residues in all WD motifs, which form the structural triad (Neer *et al.* 1994). Thus, the non-conservative mutation S301F is thought to perturb at least the structure of β-sheet 5 (WD5) and possibly the global structure of the whole protein. Consistent with this, the WD region of this mutant protein is much more trypsin sensitive than wild type ESC. The esc^{21} mimic mutation S255F was introduced into the FIE bait and surprisingly, it was found that the mutation only partially disrupted its interaction with MEA in yeast (figure 3.5).

3.4 Interaction of FIS2 and MEA

The similarity of *fis2* mutants to *fie* and *mea* suggested that FIS2 might interact with either FIE, MEA, or both. I therefore expressed full-length FIS2 as 'bait' and 'prey' fusions and tested these in yeast two-hybrid assays with FIE or MEA constructs. However, no interaction with either FIE or MEA was found (data not shown). Other groups have also reported subsequently that FIS2 does not show an interaction with either FIE or MEA in yeast (Luo *et al.* 2000 and Yadegari *et al.* 2000). One possibility is that FIS2 may only interact with FIE and MEA when the two proteins are associated in a complex. This would be analogous to the interactions between certain MADS-box containing proteins. For example, AG can interact with AP3/PI in yeast, but only when another MADS protein SEP3 is associated with AP3 or PI (Honma and Goto 2001). To test whether FIS2 could interact with a FIE/MEA complex a yeast 3-hybrid assay was carried out. MEA was expressed on a third plasmid (pTFT) together with FIE-DB and FIS2-TA in yeast host strain YB2. Again no interaction was found (data not shown). This might be attributed to steric hindrance of the full-length protein, whereby the interaction domain of the protein is restricted or obstructed in the chimeric fusion protein. Alternatively, the full-length FIS2 protein may be poorly expressed in yeast. The latter possibility is supported by the observation that both

VRN2 and SuZ12, two homologues of FIS2, cannot be expressed as full-length proteins in *E.coli* (C.Dean, personal communication). Later, A. Bishopp (ICMB, Edinburgh) showed that the EMF2 protein, which is a homologue of FIS2, does not interact in yeast 2-hybrid with the MEA homologue CLF when it is expressed as a full-length protein. However, a truncation containing only the C-terminus of EMF2, which contains the conserved VEFS box region did interact (A.Bishopp, unpublished). I therefore created FIS2VEFS-DB, which expresses the C-terminus of FIS2 as a DB fusion, and tested this in yeast two-hybrid assays with full-length MEA-TA and MEA^{4-109}–TA (figure 3.6). In this case the yeast strain AH109 was used which has a more stringent *HIS3* reporter gene than YBP2 and has a highly stringent *ADE3* reporter gene, both of which are transcriptionally dependent on reconstituted GAL4 activity. These results show that MEA and the FIS2 VEFS box can interact in yeast, and that the FIS2 interaction domain is separate from the FIE interaction domain.

3.5 Conservation of Pc-G interactions between plants and animals

It has also been shown that an amino-terminal portion of the *Drosophila* E(Z) protein is sufficient for its interaction with ESC (Tie *et al.* 1998, Jones *et al.* 1998, and Shao *et al.* 1999). This suggests that the interaction between E(Z) and ESC homologues is conserved between plants and animals, and is mediated by amino-terminal regions of E(Z) and MEA. However, there is little primary sequence similarity between these regions of MEA and E(Z) (figure 3.7), suggesting that the domain mediating the interaction with FIE/ESC has undergone sequence divergence in the plant and animal lineages. To test whether their secondary structures might be conserved in spite of the sequence divergence, I tested whether plant MEA was able to interact with *Drosophila* ESC by using the LEXA based yeast two-hybrid system (Golemis *et al.* 1996) in which the interaction of E(Z) and ESC had previously been demonstrated (Jones *et al.* 1999). TA–E(Z) and TA–MEA preys both interact with DB–ESC bait and activate the reporter genes, although the interaction with

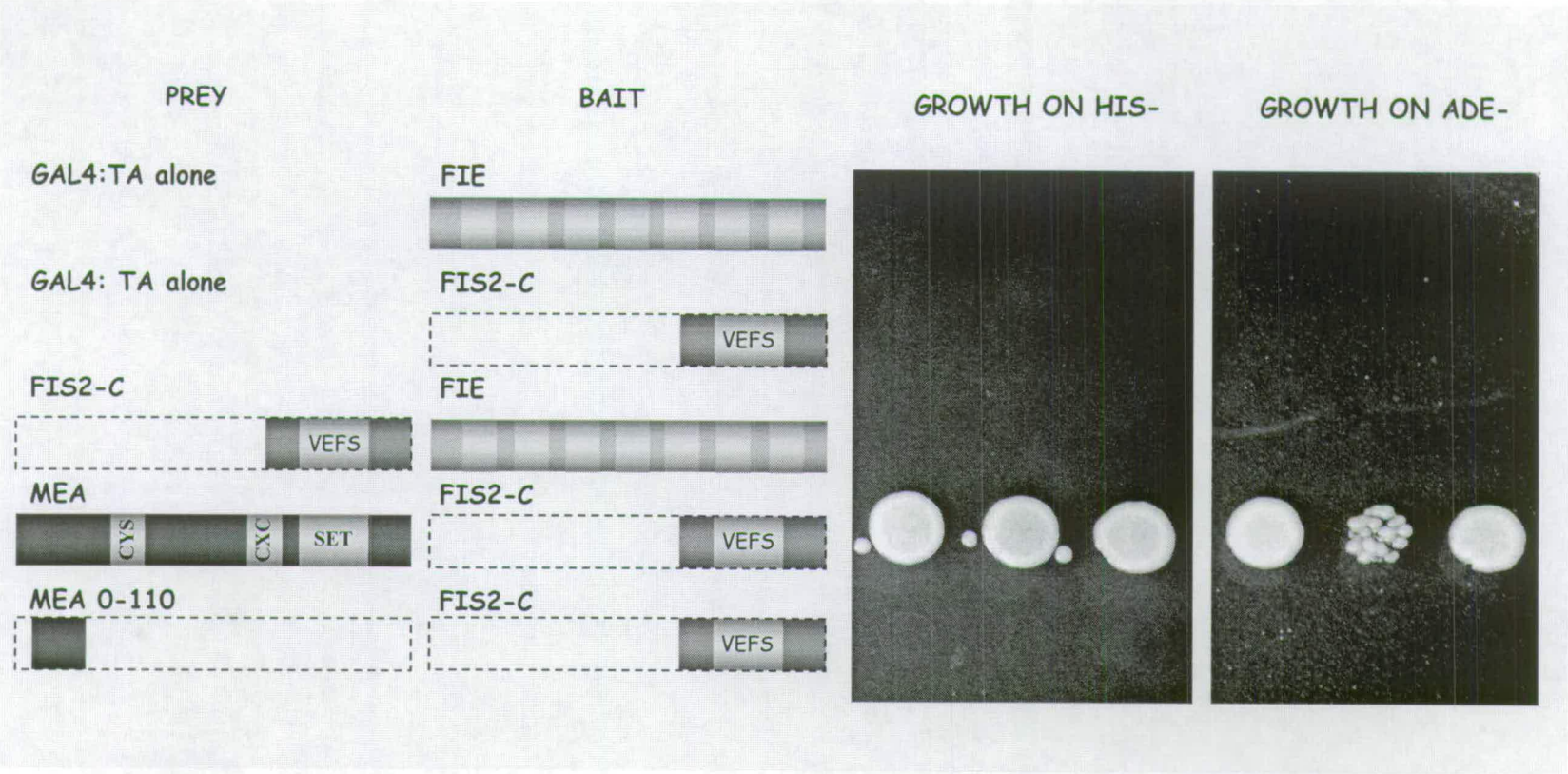


Figure 3.6 MEA Interacts with the FIS2 Carboxy-terminus in yeast. HIS and ADE reporter constructs are transcriptionally dependent on reconstituted GAL-4. Three independent transformants are shown for each growth assay on His^- and Ade^- media.

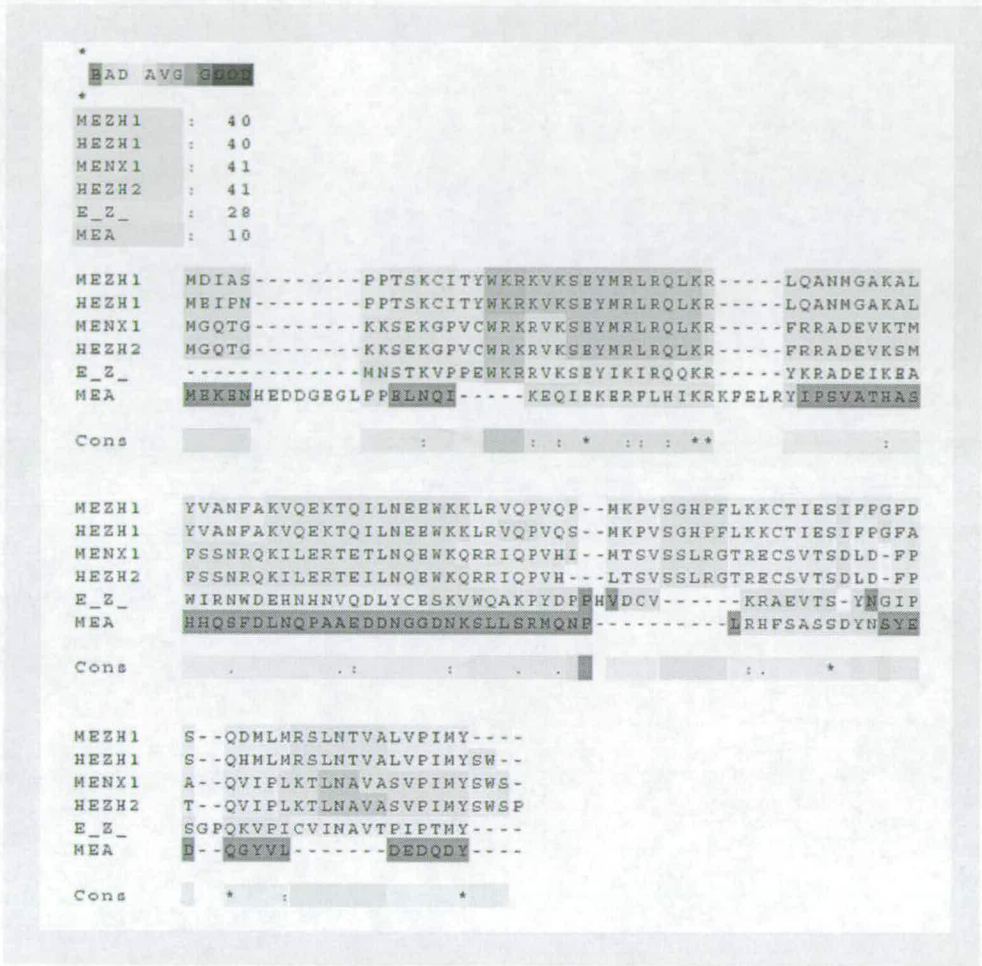


Figure 3.7 Protein alignment of the ESC/FIE binding region of E(Z)/MEA and four mammalian E(Z) homologues. The E(z) N- terminal region is highly conserved between the *Drosophila* protein and its mammalian homologues (mENX1, mEzh1, hEZH1, and hEZH2) but shares little homology with the *Arabidopsis* protein MEA. Warmer colours indicate higher consistency between the different pairwise alignments.

MEA is weaker than with E(Z) (figure 3.4). Thus MEA and E(Z) can both interact with ESC despite their divergent sequences in the amino-terminal region, suggesting they may retain a similar protein structure.

3.6 Interaction of FIE and MEA *in planta*

If the physical interaction of *FIE* and *MEA* is biologically relevant, their products would be expected to be associated together in a complex *in planta*. To test this we used fluorescence resonance energy transfer (FRET). FRET is a phenomenom that occurs when two fluorophores are in sufficient proximity (<100Å). An excited donor fluorophore *e.g.* cyan fluorescent protein (CFP) can then transfer its energy to a second acceptor fluorophore *e.g.* yellow fluorescent protein (YFP) to produce light emissions. Protein-protein interactions can be measured by fusing the proteins of interest to the flourophores and measuring the ratio between donor and acceptor emission intensities. Full-length MEA and FIE cDNAs were subcloned into the FRET vectors (pMON-YFP and pMON-CFP) to allow expression of MEA and FIE as amino-terminal fusions to YFP or CFP under the control of the CaMV35s promoter. I produced the FIE constructs and the FRET analysis was subsequently carried out by C. Spillane (IPB, University of Zürich, Switzerland) as part of a collaboration. Unfortunately no FRET was observed with either combinational pair of the fusion proteins (MEA-CFP/FIE-YFP and MEA-YFP/FIE-CFP) when expressed in cowpea protoplasts. However, this result does not exclude the possibility that FIE and MEA interact *in planta,* as FRET may not occur for several reasons. For example, due to steric hindrance of at least one of the fusion protein pair, whereby the interaction domain of the protein is restricted or obstructed in the fused protein. Alternatively, the fluorophores may not be juxtaposed sufficiently for FRET to occur.

Cotransfections of cowpea protoplast cells with *FIE–CFP* (figure 3.8a) and *MEA–YFP* (figure 3.8b) indicate that FIE and MEA co-localize to the nucleus and are both consistently excluded from the nucleolus. Nuclear localization of both MEA and FIE was also observed in single transfections and for the reciprocal constructs FIE–YFP (figure 3.9) and MEA–CFP (data not shown). The *in vivo* co-localization of MEA and FIE proteins in the nucleus of plant cells, in addition to the demonstration of their physical interactions in yeast (figures 3.3 and 3.4) and *in vitro*, is consistent with a MEA–FIE Pc-G protein complex occurring *in vivo*.

3.7 Genetic interaction of *FIE* with *CLF*

Unlike MEA, FIE is a single copy gene and is expressed after seed development during vegetative growth. This suggested that FIE could be a partner for the other plant E(z) members, CLF and CLK (EZA1). Consistent with this, CLF and CLK also interact with FIE in yeast (C. MacDougall and Y. Chanvivattana, unpublished). To determine if *FIE* and *CLF* interact genetically *fie-2/+ clf-2/clf-2* double mutant plants were generated. The *fie-2* allele was introduced from a male parent in crosses to *clf-2* plants since *fie* alleles cannot be transmitted maternally. If the female gametophyte inherits the *fie-2* allele the resultant embryo aborts; therefore *fie-2* homozygous plants could not be generated. However, a reduced dosage of *FIE* might have an effect in a genetic background that is already compromised for a different member of the complex, i.e. *CLF*. Figure 3.10 compares the vegetative phenotype of *clf-2/clf-2* plants with one or two FIE^{+} alleles, and shows that there is no effect of FIE^{+} dosage. The developing seeds of *clf-2/clf-2 fie-2/+* plants were also examined but no enhancement or suppression of the *fie* phenotype was seen (data not shown). However, these results do not exclude a genetic interaction between *FIE* and *CLF*, as members of a common complex do not always show dosage effects. One example is the complex containing the B-function homeotic proteins PI and AP3. Here, *pi/+ ap3/+* trans-

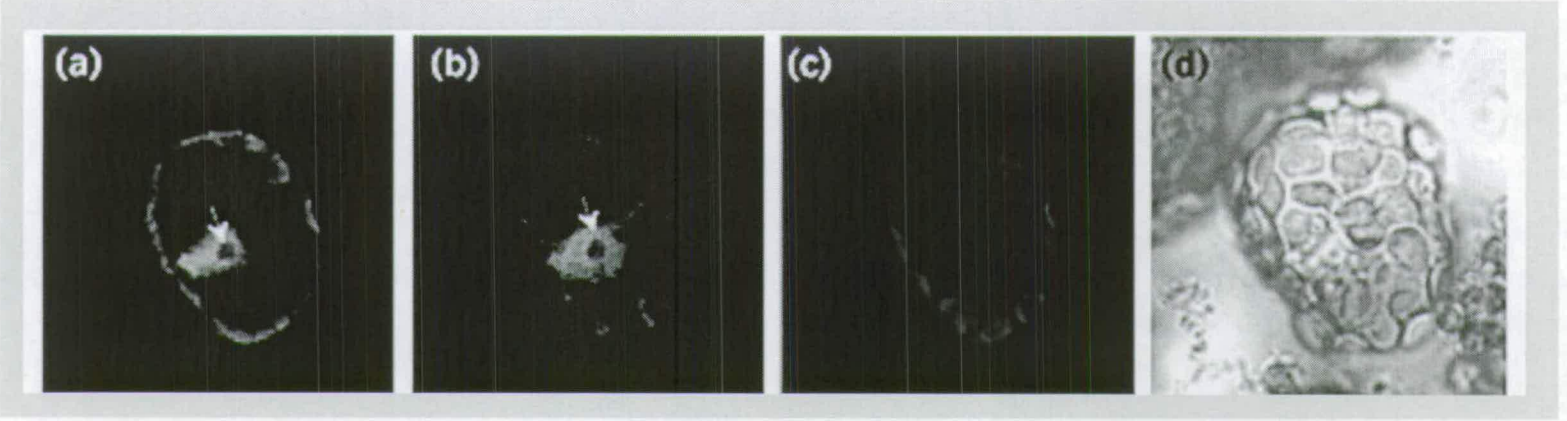


Figure 3.8 Co-localisation of FIE-CFP and MEA-YFP Fusion Proteins. Cowpea protoplast co-transfected with pMON-FIE-CFP and pMON-MEA-YFP. (a) FIE-CFP, (b) MEA-YFP, (c) chloroplast autofluorescence, (d) phase contrast microscopy. White arrows indicate the nucleolus. Co-localised FIE-CFP and MEA-YFP in the cytoplasm may be related to high levels of expression from the CaMV 35S promoter.

Figure taken from Spillane *et al.*, 2000.

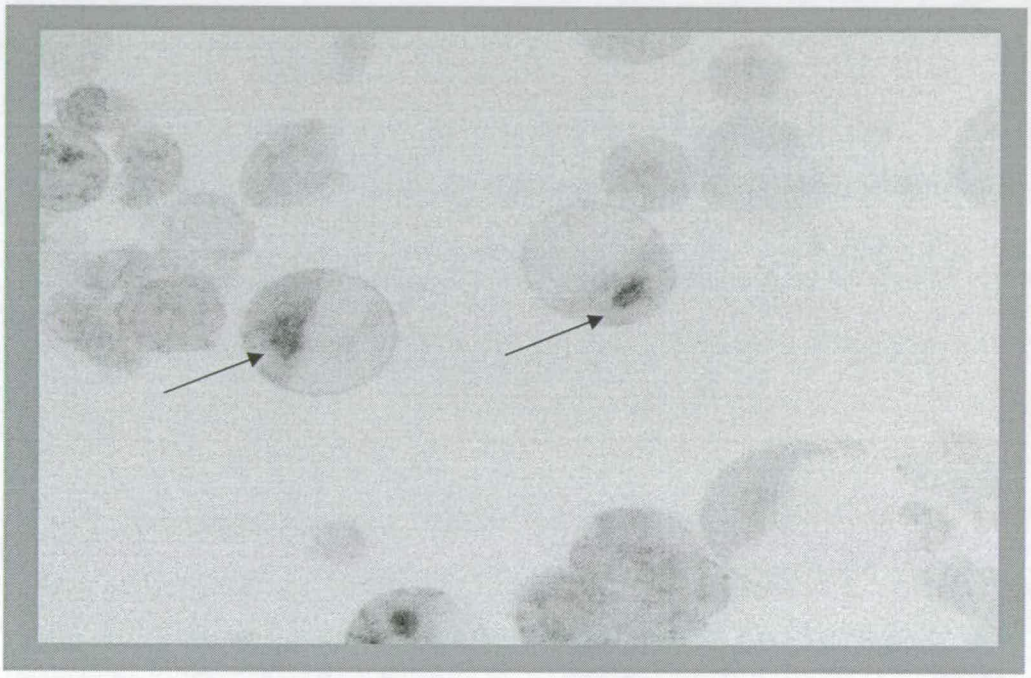


Figure 3.9 YFP-FIE Fusion Proteins are Nuclear Localised in Cowpea Protoplasts. Cowpea protoplasts transfected with pMON-FIE-YFP. Black arrows indicate nucleolus. (Image supplied by C. Spillane).

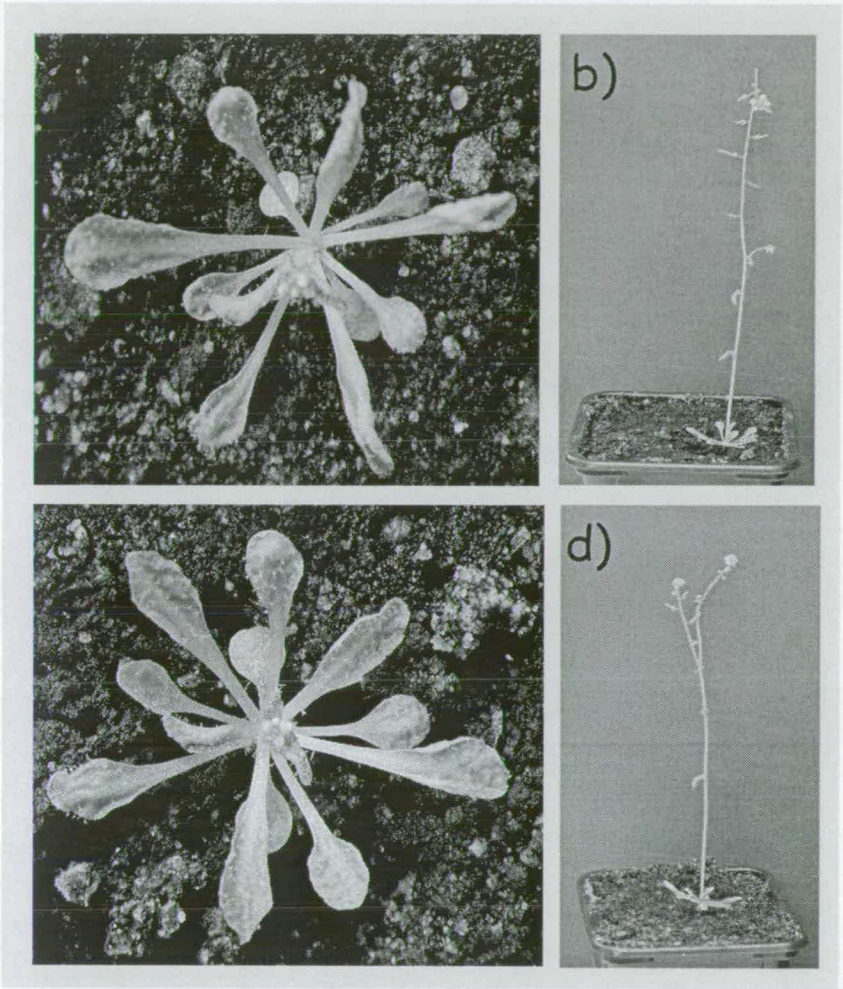


Figure 3.10 Altered dosage of FIE in *clf* background. a) *clf-2* homozygote three weeks after germination. b) *clf-2* homozygote four weeks after germination. c) *clf-2/clf-2 fie-3/+* three weeks after germination. d) *clf-2/clf-2 fie-3/+* four weeks after germination. No effect of *FIE* dosage on *clf* phenotype is seen.

heterozygotes appear wild type, and *ap3/+ pi/pi* plants do not have a more severe phenotype than *+/+ pi/pi* plants (Bowman *et al.* 1989).

3.8 Introduction: Novel FIE interactions

The FIS proteins (FIE, MEA and FIS2) act together in a complex presumably to repress target genes. The *PHERES1* (*PHE1*) gene encoding a MADS protein was recently identified as a target for FIE and MEA repression (Köhler *et al.* 2003). However, as none of the FIS proteins have been found to possess specific DNA-binding activity it remains unclear how this complex could be targeted to its target gene(s). It is known that one FIS complex is 600kDa, and it has recently been shown that MSI1 is a member of this complex. The MEA:FIE:MSI1 component of this complex is 169kDa (Köhler *et al.* 2003) and FIS2 has a molecular weight of 78 kDa. As these 4 proteins have not been shown to form homomeric dimers it suggests that other unidentified proteins might interact with the 3 FIS and MSI1 proteins. The complex could contain proteins that act to target the complex to its DNA target gene(s) (*e.g. PHE1*). To identify factors interacting with FIE that could function in this way, a yeast two-hybrid screen was carried out. FIE-DB was used as 'bait' and screened against a TA library constructed from *Arabidopsis* cDNA from roots, seedlings, flowers and siliques (a generous gift of Dr W.Crosby).

3.9 Mapping the FIE interaction domain of RAP-1

Of the $2x10^6$ transformants screened, 24 putative FIE interacting proteins activated the *HIS3* reporter gene but only one of these activated both the *HIS3* and *lacZ* reporter genes. The prey plasmid was isolated from this colony and sequenced and the interacting protein was identified as R-homologous *Arabidopsis* Protein-1 (RAP-1), a member of the basic helix-loop-helix (bHLH) superfamily of myc-like transcription factors (figure 3.11a). In order to map the region of RAP-1 sufficient for its interaction with FIE a series of TA-RAP-1 deletion constructs were generated based on existing restriction sites within the RAP-1 cDNA (figure 2.2). These constructs were then co-transformed with DB-FIE into yeast host strain AH109 for two-hybrid analysis. Figure 3.12 shows that amino acids 234 to 353 show

border primer LBb1 used to confirm the T-DNA insertion was prone to mispriming in standard PCR conditions, giving a band of 450bp (yellow arrow in figure 3.13a). However, increasing the annealing temperature in the PCR programme to 62^{0}C enabled the correct priming of this oligonucleotide. Figure 3.13a shows the identification of an individual homozygous for the T-DNA insertion in the line salk_017005. The fragment amplified was cloned and sequenced. The T-DNA insertion was found to be at nucleotide +1025 (relative to the ATG) as annotated on the SIGnAL website. Figure 3.13b shows the identification of individuals heterozygous for the T-DNA insertion in the line salk_061267. In this case the T-DNA specific primer 'salk border' was used instead of LBb1. The insertion band was cloned and sequenced. The T-DNA insertion was found to be at nucleotide +935 (relative to the ATG), slightly upstream of that described on the SIGnAL website. The insert at +1025 was designated *rap-1-1*, and the insert at +935, *rap-1-2*. Both insertions are predicted to cause severe loss-of-function or null mutations, as they are upstream of the highly conserved bHLH domain that is probably essential for RAP-1 function. The T-DNA insertions were also confirmed using Southern blot hybridisation (figures 3.14-3.15). Genomic DNA was extracted and digested with the restriction enzyme *Bgl*II. The high proportion of *rap-1-2* homozygotes shown in figure 3.14 was probably due to the fact that most of the plants used for DNA analysis had been previously selected for strong kanamycin resistance, which may have enriched for those carrying two T-DNA copies, as kanamycin resistance was weak in this line.

Homozygous *rap1-1* and *rap1-2* plants (confirmed by PCR) were examined for mutant phenotype. If the RAP-1 interaction with FIE observed in yeast two-hybrid is biologically relevant one might expect to see a *fie*-like mutant phenotype such as seed abortion or autonomous endosperm development. However, no difference from wild type was observed (figure 3.16). Because FIE is likely to act with other Pc-G members after germination,

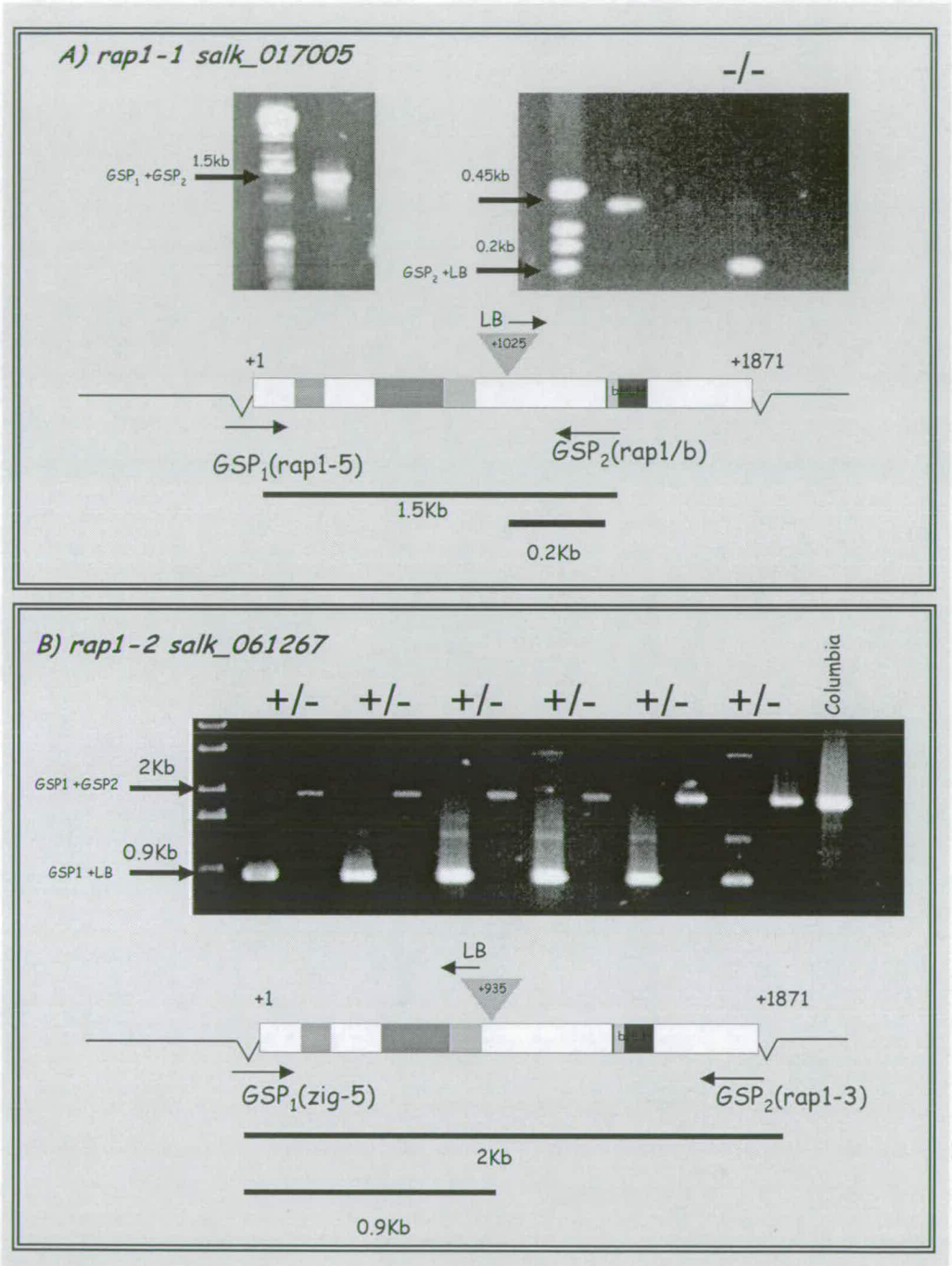


Figure 13.13 Identification of *RAP-1* T-DNA insertions by PCR. A) shows PCR of *rap1-1 salk_017005* T-DNA insertion line. Red arrow indicates wild type band. Green arrow indicates insertion band. Yellow arrow indicates mis-primed artefact band. The insertion was confirmed at +1025 by sequence analysis. B) Shows PCR of single plants from *rap1-2 salk_061267* insertion line. For each individual, two PCR reactions are shown, the first using GSP_1 and LB to identify insert allele, the second using GSP_1 and GSP_2 to identify RAP-1$^+$ allele. Red arrow indicates wild type band. Green arrow indicates insertion band. The insertion was confirmed at +935 by sequence analysis.

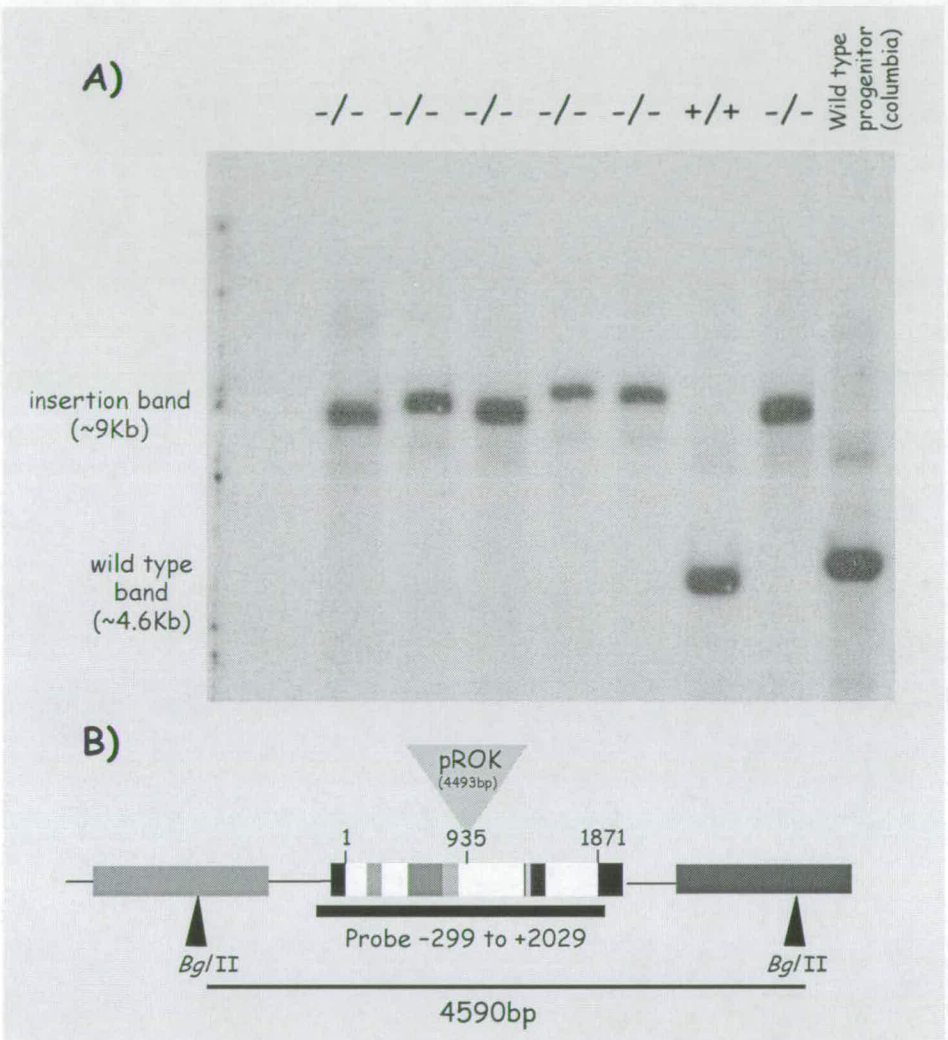


Figure 3.14. Southern blot analysis of *RAP-1* insertion lines. A) Genomic DNA from single plants from the Salk_061267 family was cut with *Bgl* II and probed on a Southern blot with *RAP-1* sequence. The far right lane shows the wild type progenitor (Columbia) band. A T-DNA insertion is identified by a band shift. -/-, homozygous for the insertion at the *RAP-1* locus, +/+, wild-type *RAP-1* locus. B) Diagram of *RAP-1* locus showing position of T-DNA insertion and probe.

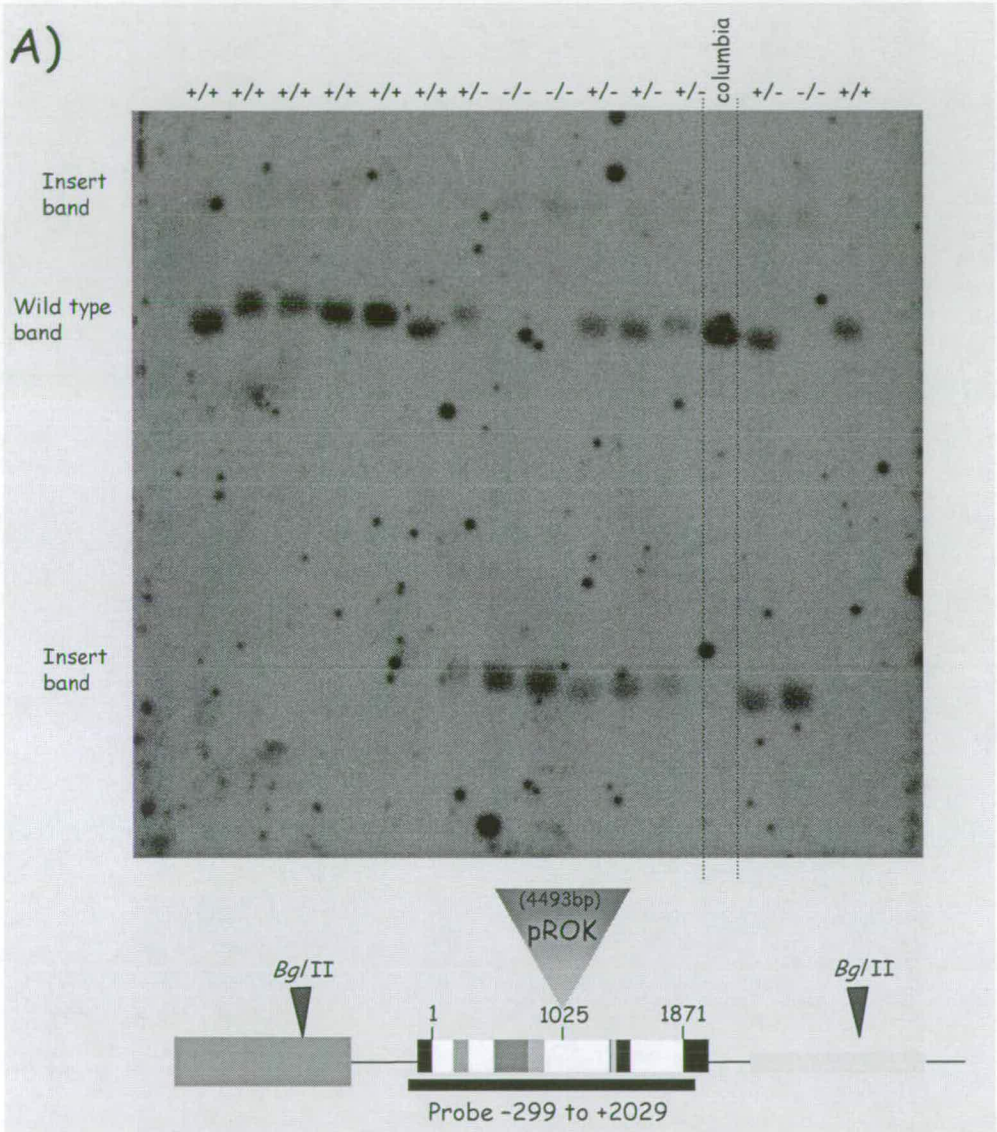


Figure 3.15 Southern blot analysis of *RAP-1* insertion lines. A) Genomic DNA from single plants from the salk_017005 family was cut with *Bgl* II and probed on a Southern blot with *RAP-1* sequence. The far right lane shows the wild type progenitor (Columbia) band. A T-DNA insertion is identified by a band shift. -/-, homozygous for the insertion at the *RAP-1* locus, +/+, wild-type *RAP-1* locus. B) Diagram of RAP-1 locus showing position of T-DNA insertion and probe.

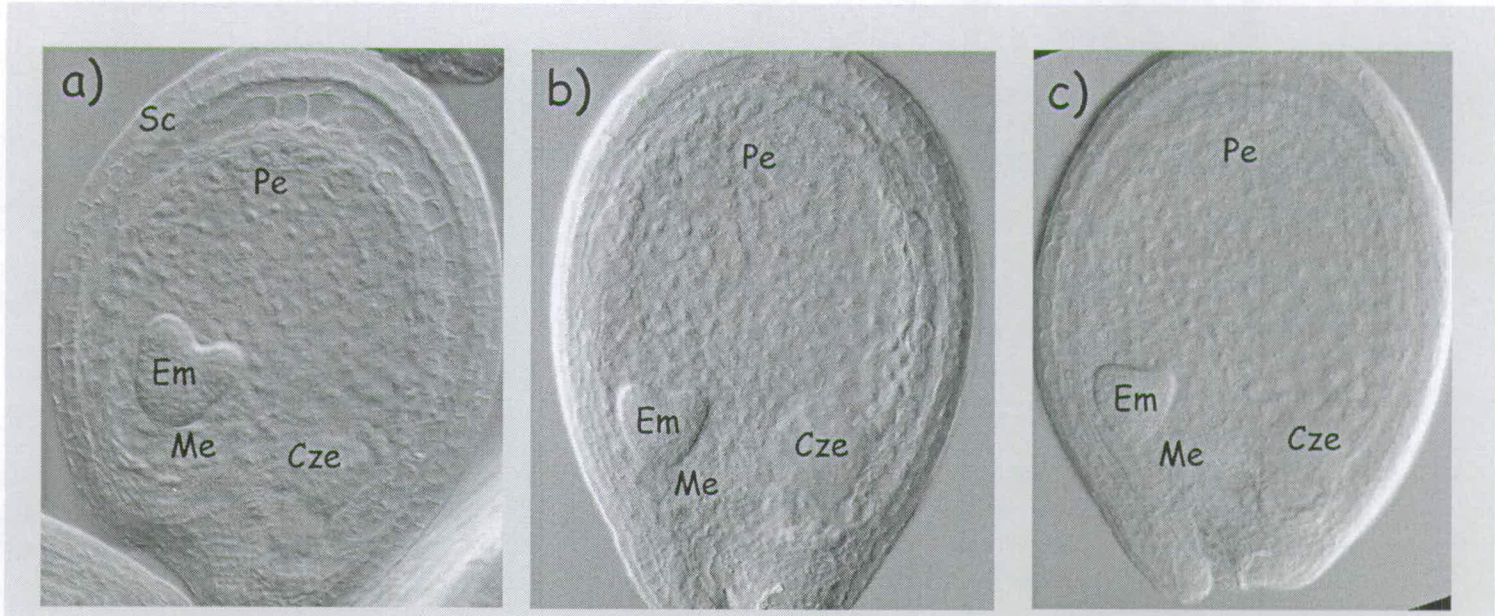


Figure 3. 16 Developing Seed of wild type and *rap-1* mutants. DIC light microscopy of cleared developing seed. a) Wild type b) *rap-1-1/rap-1-1,* c) *rap-1-2/rap-1-2.* Em – Embryo, Cze – Chalazal endosperm, Sc – Seed coat, Me – micropylar endosperm, Pe – Peripheral endosperm.

whole plant phenotypes (for example, similarities to *clf* mutants) were also characterised. However, no apparent vegetative or floral mutant phenotype was observed (data not shown).

3.11 *RAP-1* redundancy

The *Arabidopisis* genome has undergone much duplication and gene rearrangement during evolution (McGrath *et al.* 1993, Kowalski *et al.* 1994, Paterson *et al.* 1996, Lin *et al.* 1999, Mayer *et al.* 1999, Terryn *et al.* 1999, Blanc *et al.* 2000). Analysis of the completed genome sequences revealed many tandem gene duplications as well as larger scale segmental duplications on different chromosomes. This has led to the suggestion that *Arabidopsis* is descended from an ancient tetraploid, and has undergone whole genome duplication early on in the *Arabidopsis* lineage (Blanc *et al.* 2000). Consequently, the *Arabidopsis* genome is likely to contain a high proportion of functionally redundant genes. Numerous examples abound, for example the genes *SEP1*, *SEP2* and *SEP3* encode similar MADS-box transcription factors. These genes show a similar expression pattern in flowers, and behave redundantly so that single and double mutants show no phenotype (Pelaz *et al.* 2000). To identify which members of the bHLH family possess high primary sequence homology with RAP-1 and hence might share its biological function, a BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*1990) search was carried out on the *Arabidopsis* protein sequence database using the full-length RAP-1 protein sequence, and a phylogenetic tree was generated using CLUSTAL X (figure 3.17). bHLH5/ATR2, bHLH28, and bHLH4/AtMYC4 were found to cluster in a sub-clade with RAP-1. The SIGnAL database at the time only contained T-DNA inserts in two of the genes: bHLH5/ATR2 and bHLH28. These two insertion lines were ordered and processed in the same way as previously described for *rap-1-1* and *rap-1-2* to identify plants carrying T-DNA insertions (section 3.10). Unfortunately, PCR and Southern analysis showed that none of the 11 plants tested for line SALK_059620 carried the predicted insertion at bHLH28 (figure 3.18), indicating either that the Salk

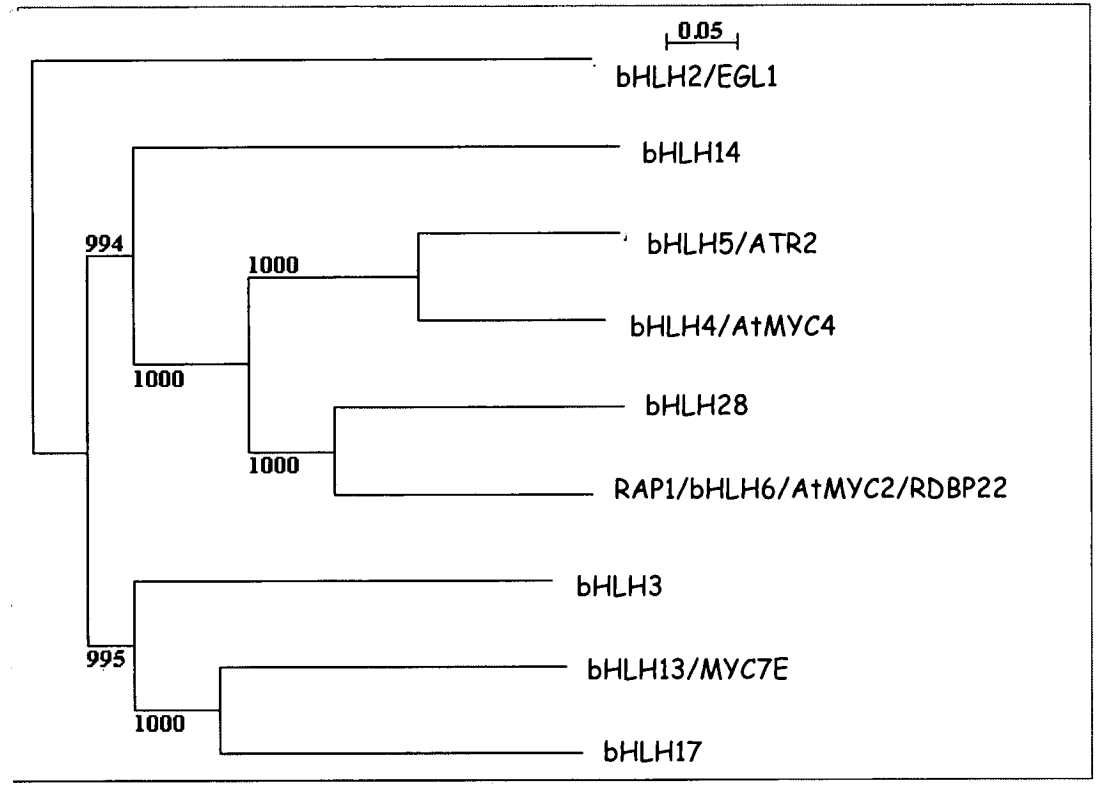


Figure 3.17 Phylogenetic tree of bHLH family closest to RAP-1. The unrooted tree, constructed using Clustal W, shows the evolutionary relationships between RAP-1 (red) and its closest homologues. The tree was constructed using the entire length amino acid sequence of each protein. Numbers represent bootstrap values.

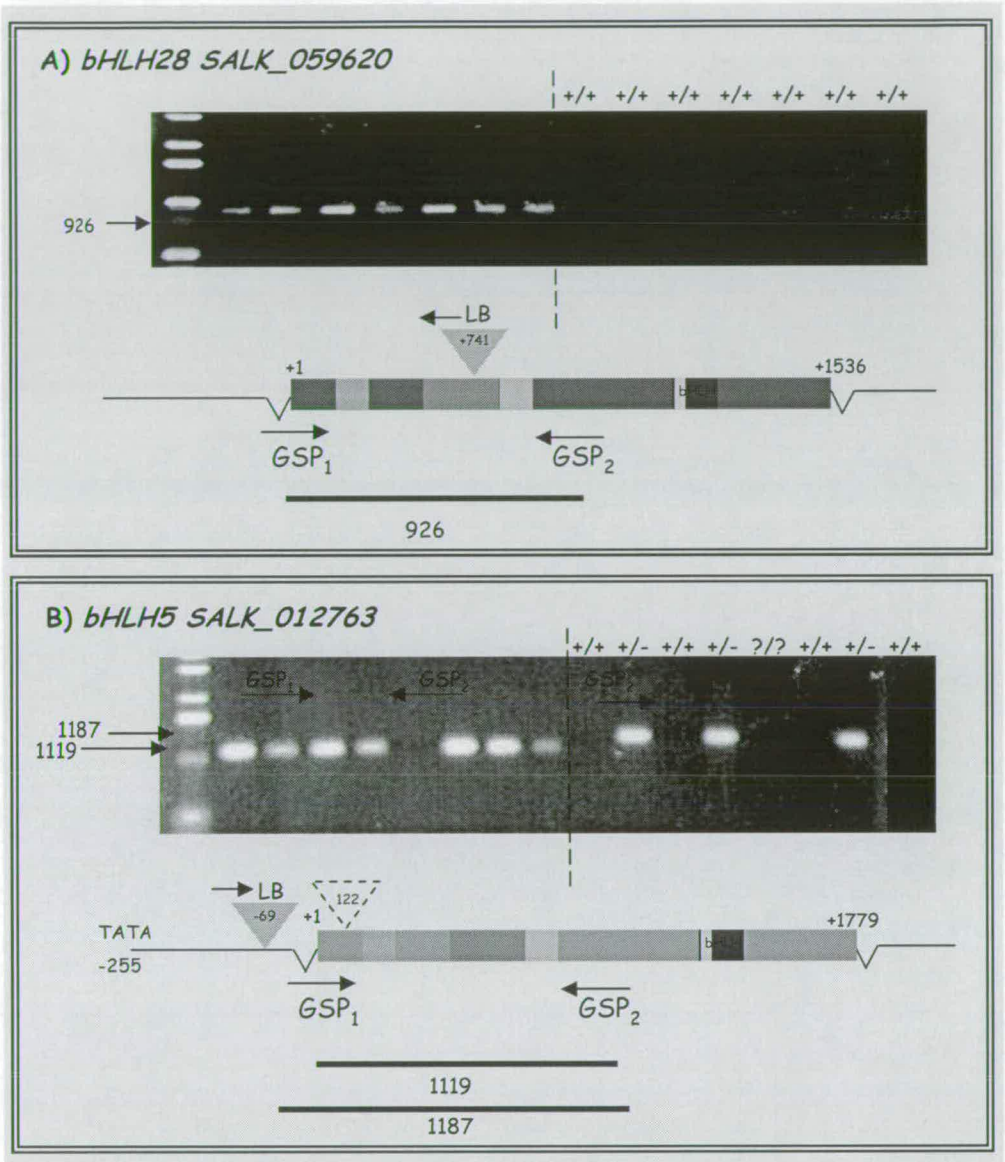


Figure 13.18 PCR analysis of *Salk* 059620 and 012763 family. A) shows PCR of DNA from 7 individual plants from *bHLH28 Salk_059620* T-DNA insertion line. First 7 lanes show reactions using GSP_1 and GSP_2 to identify $bHLH28^+$ allele. The second 7 lanes show reactions using GSP_1 and LB to identify insert allele. Red arrow indicates wild type band. No insertion was found. B) Shows PCR of DNA from 8 individuals from *bHLH5 Salk_012763* insertion line. First 8 lanes show reactions using GSP_1 and GSP_2 to identify $bHLH5^+$ allele. The second 8 lanes show reactions using GSP_2 and LB to identify insert allele Red arrow indicates wild type band. Green arrow indicates insertion band. The insertion was confirmed at -69 by sequence analysis.

database has been wrongly annotated or that plants carrying the insertion were rare in this family and were not present in the small sample analysed. An insertion was found at the bHLH5/ATR2 locus (figure 3.18) but was upstream of the ATG start. The SIGnAL database describes the insertion at +122 (relative to the ATG) but it was found to be at –69. The allele was designated '*bhlh5-1*'. It is thought that this insertion could still be mutagenic as it could interfere with promoter binding or with translation in the 5'UTR.

Homozygous *bhlh5-1* plants (confirmed by PCR) have a wild type phenotype (data not shown), and the developing seed display no sign of a mutant phenotype (figure 3.19). No loss-of-function alleles have been previously described, although one dominant allele has been characterized. The dominant *atr2D* mutation creates an aspartate to asparagine change at a position near the amino terminus of the protein (Smolen *et al.* 2002). *atr2D* plants display increased expression of several tryptophan genes as well as a subset of other stress related genes. However, overexpression of the mutant *atr2D* cDNA but not the wild type *ATR2* gene in transgenic plants conferred pleiotropic effects, including reduced size, dark pigmentation, and sterility. Therefore, it was concluded that *atr2D* is likely to be an altered function allele and the wild type allele may not therefore be involved in tryptophan metabolism (Smolen *et al.* 2002).

bhlh5-1 plants were crossed to *rap-1* plants to produce the double mutant, so that if there was functional redundancy between the two genes, the double mutant might reveal a loss-of-function phenotype. However, it was found that plants homozygous for both mutant alleles (confirmed by PCR) do not show vegetative (data not shown) or developing seed (figure 3.19) mutant phenotypes.

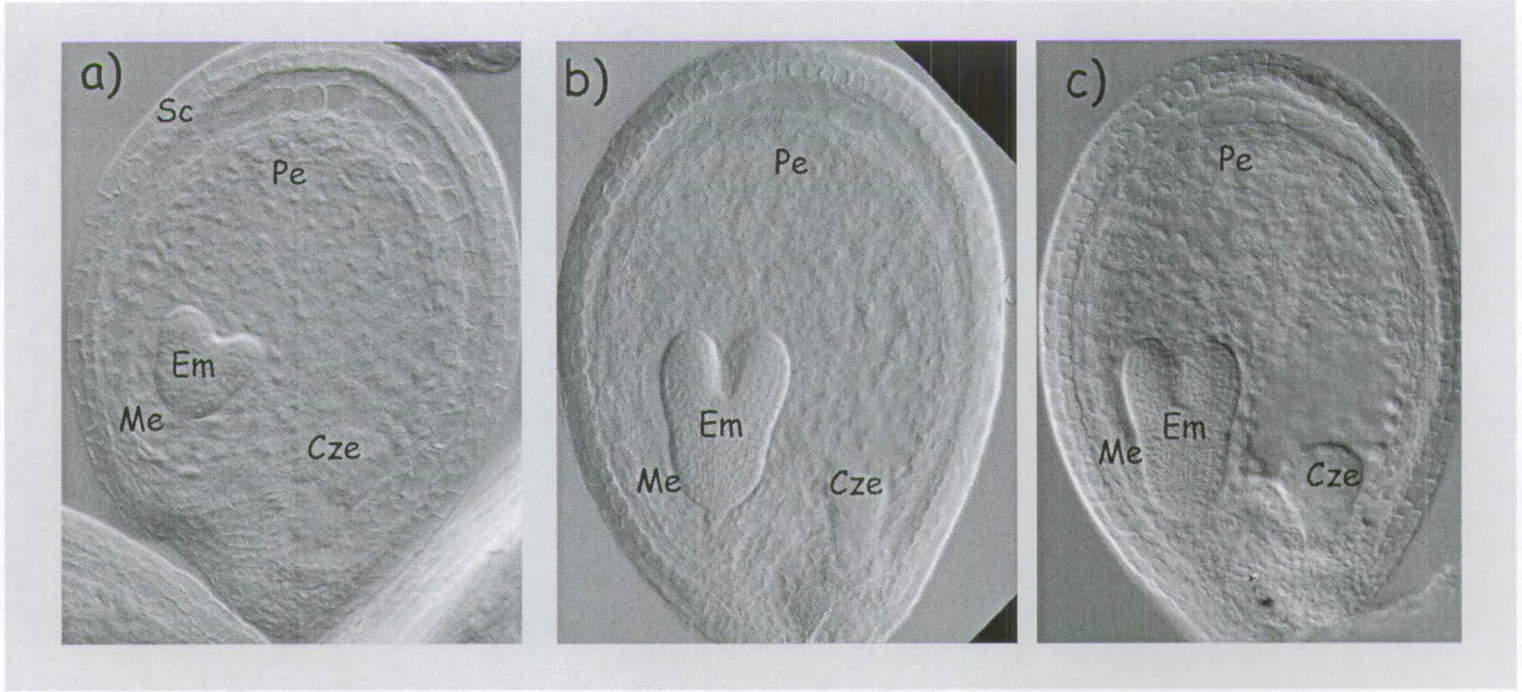


Figure 3. 19 Developing Seed of wild type and *bhlh5* mutants. DIC light microscopy of cleared developing seed. a) Wild type, b) *bhlh5/bhlh5, c) bhlh5/bhlh5, rap-1-1/rap-1-1.* Em – Embryo, Cze – Chalazal endosperm, Sc – Seed coat, Me – micropylar endosperm, Pe – Peripheral endosperm.

3.12 Genetic interactions with *fie*

To further test for a genetic interaction of *RAP-1* or *bHLH5/ATR2* with *FIE in planta*, the mutant alleles were introduced into a *fie* background to investigate whether the *fie* phenotype was modified. Plants heterozygous for *fie* produce 50% seed that display characteristic abnormalities as described previously (section 1.4). If RAP-1 or bHLH5 acts with FIE in a complex one might expect to see a genetic interaction, for example enhancement of seed or vegetative phenotype. To determine if *FIE* and *RAP-1/bHLH5* interact genetically *fie-2/+ rap-1-1/rap-1-1, fie-2/+ rap-1-2/rap-1-2, fie-2/+ bhlh5-1/bhlh5-1* double mutant plants were generated. The *fie-2* allele was introduced from a male parent in the crosses since *fie* alleles cannot be transmitted maternally. If the female gametophyte inherits the *fie-2* allele the resultant embryo aborts; therefore *fie-2* homozygous plants could not be generated. However, a reduced dosage of *FIE* might have an effect in a genetic background that is already compromised for a different member of the complex, *e.g. RAP-1* or *bHLH*. Figure 3.20 shows the phenotype of *fie* seed in *rap-1* and *bhlh5-1* backgrounds. Again, no differences from the characteristic *fie* phenotype were observed: 50% of the seed aborted with a misshapen embryo and abnormal endosperm as seen in *fie* single mutant seed. The vegetative phenotype of these plants was also examined but no mutant phenotype was seen (data not shown). However, these results do not exclude a genetic interaction between *FIE* and *RAP-1* and *bHLH5* for several reasons: members of a common complex do not always show dosage effects, *RAP-1* and *bHLH5* may be redundant, and the *rap-1/bhlh5-1* phenotype may be subtle or not apparent with the growth conditions used.

FIE is a single copy gene that is also expressed after seed development during vegetative growth. This suggests that FIE could be a partner for the other plant E(z) members besides MEA. Consistent with this, CLF and CLK also interact with FIE in yeast (C. MacDougall and Y. Chanvivattana, unpublished). As RAP-1 also interacts with FIE in yeast, there is a

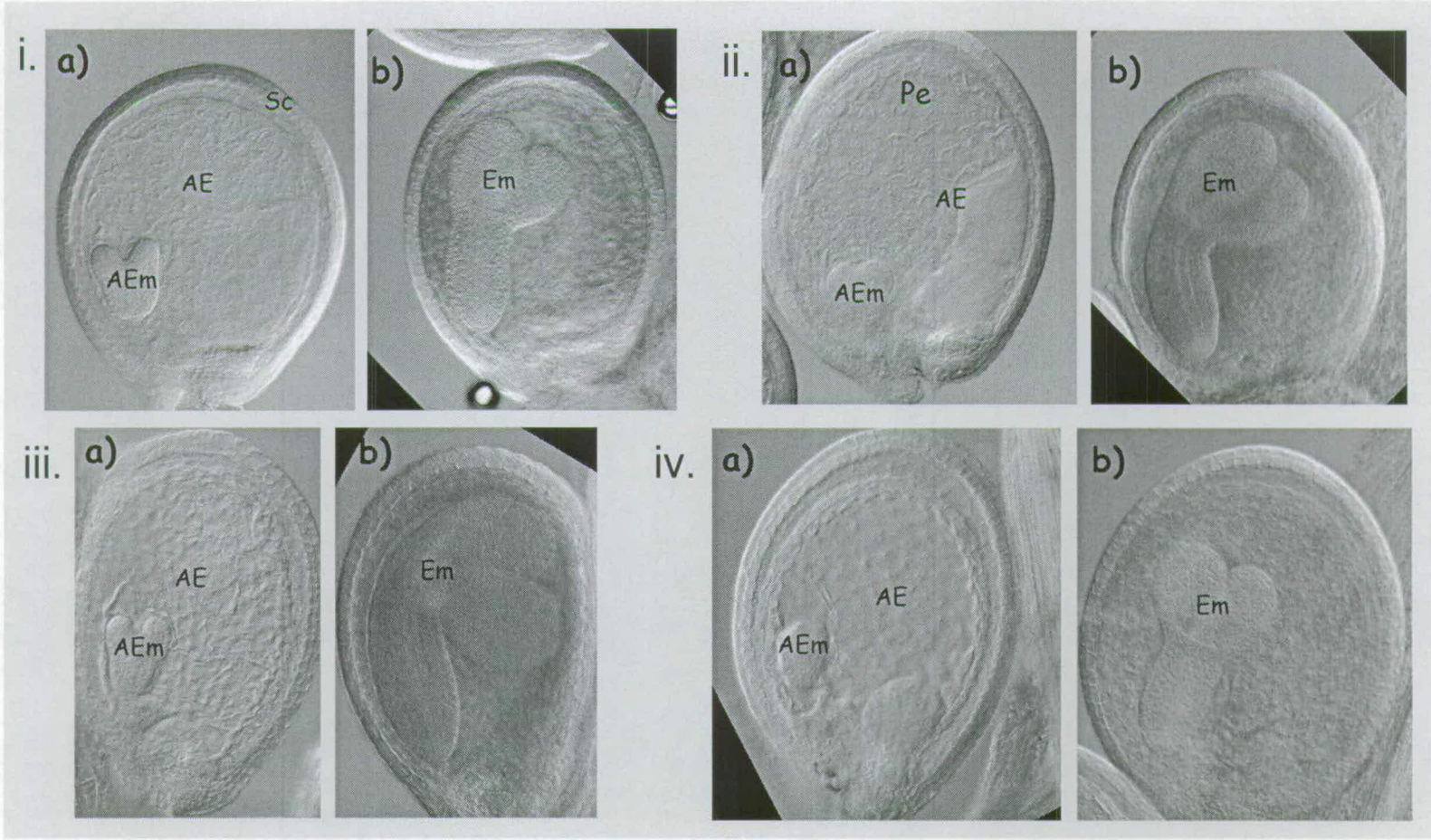


Figure 3.20 *rap-1* and *bhlh5* in a *fie* background. DIC light microscopy of cleared developing seed. a) and b) show seed at the same time after fertilization, a) shows *fie* seed showing aborting embryo and abnormal endosperm, b) shows FIE^{+} phenotype i.) *bhlh5/bhlh5 fie/+,* ii.) *rap-1-2/rap-1-2 fie/+,* iii.) *rap-1-1/rap-1-1 fie/+, iv) fie/+.* Em – Embryo, AEm- Abnormal embryo, Sc – Seed coat, AE – Abnormal Endosperm.

potential for genetic interaction between *RAP-1* and *CLF in planta*. *rap-1-1* and *rap-1-2* were therefore crossed into the *clf* background (*clf50*). One in 16 of the F2 individuals will be homozygous for both *rap-1* and *clf* and may show an enhancement of the *clf* phenotype. However, no such effect was observed (figure 3.21). In both crosses a ratio of 3:1 wild type:*clf* phenotypes was observed (*rap-1-2/clf50:* $\chi 2$=2.748 1d.f. p>0.1, *rap-1-1/clf50:* $\chi 2$=0.162 1d.f. p>0.5). The *bhlh-5-1* allele was also crossed into a *clf* background (*clf-2*), and a ratio of 3:1 wild type:*clf* was observed (70:24, $\chi 2$=0.014 1d.f. p>0.9). Figures 3.22a-c show representatives from the F2 population. Figure 3.22d shows a *clf-2/clf-2* individual of the same age. Figures 3.22e-f show the same plants 2 weeks later. Two classes of *clf* severity were observed. The weaker *clf* class (figure 3.22b) does not correspond to the expected frequency of a double mutant, 1 in 16 (observed = 70 wt: 9 severe *clf*: 15 weaker *clf*, $\chi 2$= 18.397, 2 d.f. p<0.005). However, wild type Columbia and *clf-2* crosses need to be made for comparison in order to establish if the classes of *clf* severity observed could be attributed to the mixed ecotypes of Columbia and Landsberg *erecta* in the F2 population. Also, it cannot be excluded that the weaker *clf* phenotype corresponds to a *clf-2* phenotype in a mixed background, and the stronger *clf-2* phenotype (figure 3.22c) corresponds to the double mutant (expected frequency 1 in 16) (observed = 70 wt: 15 severe *clf*: 9 weaker *clf*, expect 70.5: 17.625: 5.875, $\chi 2$=2.057, 2d.f, p>0.1). Taken together these results do not support a genetic interaction between *CLF* and *RAP-1*.

3.13 *RAP-1* expression pattern

In situ hybridisation was carried out to examine the expression pattern of *RAP-1*. A region of the *RAP-1* cDNA that is poorly conserved was identified, so as to minimise cross hybridisation of the probe with homologous genes. This region was subcloned into pBluescript SK+ and KS$^-$ and labelled RNA probe were generated by *in vitro* transcription incorporating digoxygenin-11-UTP and hybridised to sections of *Arabidopsis* inflorescence.

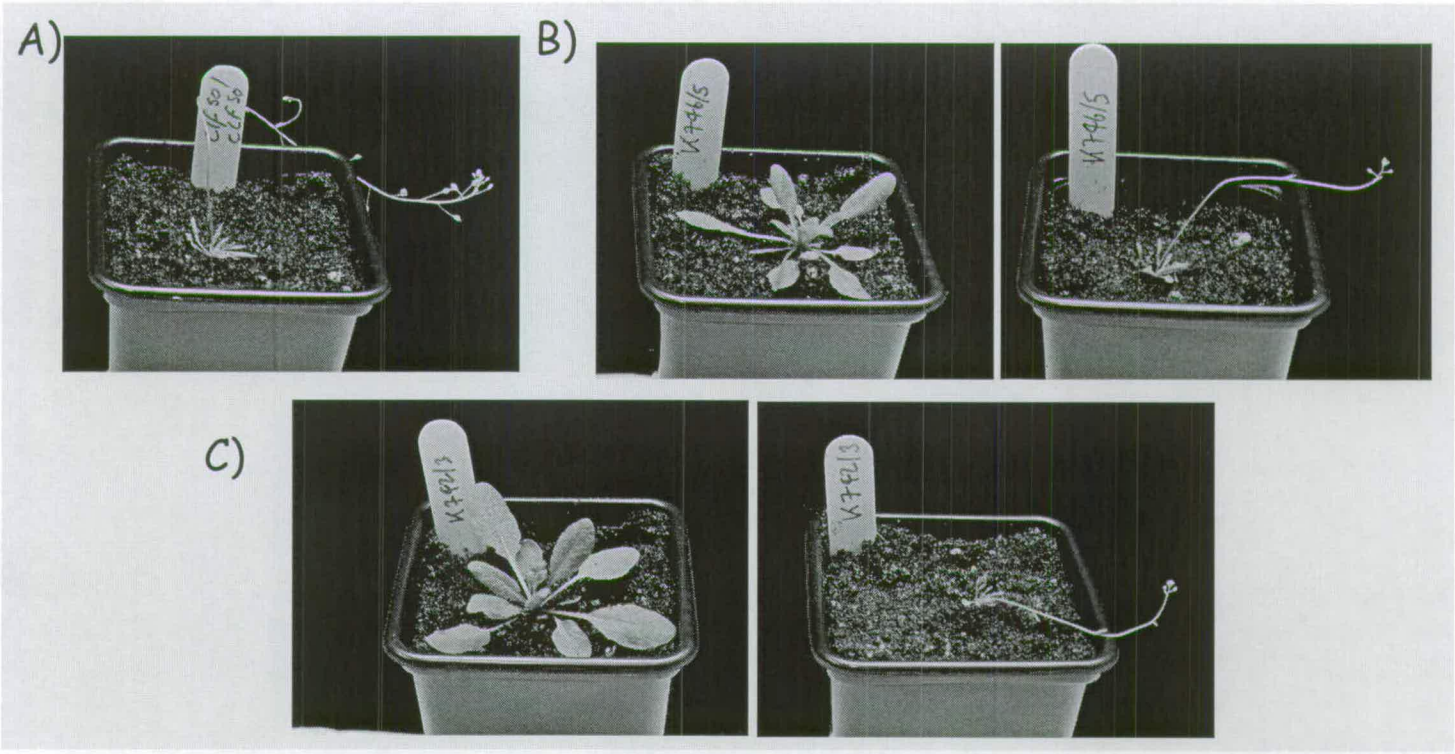


Figure 3.21 Phenotypes in F2 of *rap-1* X *clf50* cross. A) *clf50/clf50* (Ws). B) *rap-1-1* x *clf50* F2 population. i. Wild type phenotype n=233. ii. *clf* phenotype n=95. C) *rap-1-2* x *clf50* F2 population. i. Wild type phenotype n=54. ii. *clf* phenotype n=20.

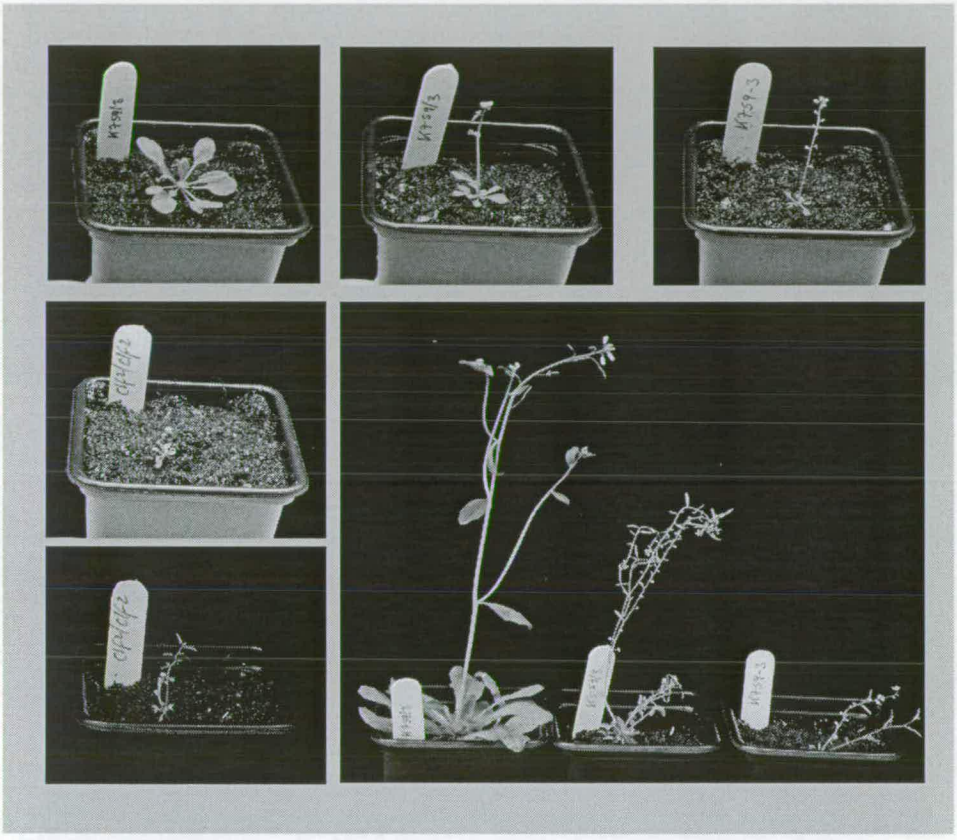


Figure 3.22 Genetic analysis between *bhlh5-1 and clf-2*. a)-c), *bhlh5-1* x *clf-2* F2 population (4 weeks old). d) *clf-2/clf-2* (Lansberg *erecta*) (4 weeks old). e) *clf-2/clf-2* (Lansberg *erecta*) (6 weeks old). f) *bhlh5-1* x *clf-2* F2 population (6 weeks old). i. n=70, ii. n=15, iii. n=9.

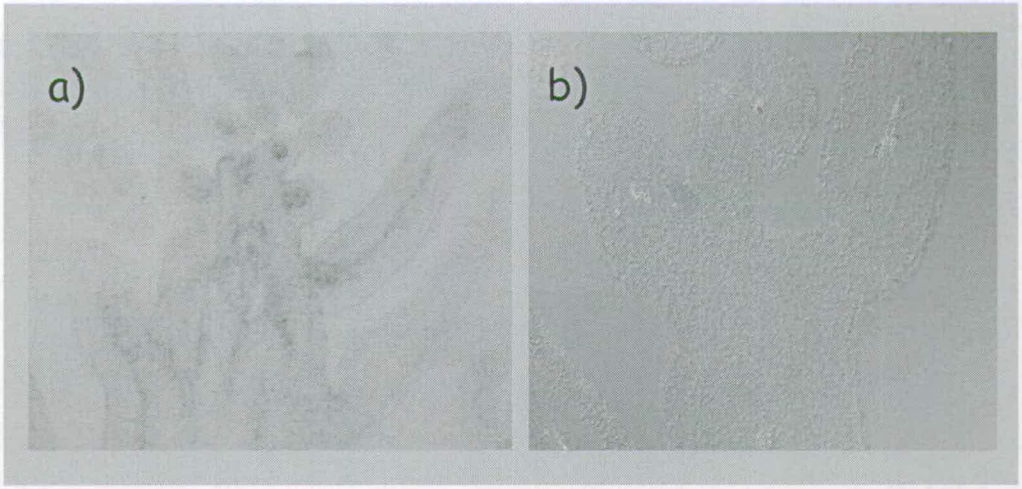


Figure 3.23. *RAP-1* in *situ* hybridisation. A RNA probe was designed to minimise cross hybridisation with homologous genes. *In vitro* transcribed anti-sense and sense RNA probe was non-radioactively labelled with digoxygenin-11-UTP and hybridised to longitudinal sections of wild type *Arabidopsis* inflorescence. a) Anti-sense. Dark staining representing mRNA can be seen as a diffuse pattern but is most intense in the vasculature. b) Sense.

Figure 3.23 shows a cross-section through the inflorescence hybridised with anti-sense *RAP-1* probe. Dark staining representing mRNA can be seen as a diffuse pattern but is most intense in the vasculature. Sense *RAP-1* probe shows no dark staining.

Chapter 4. The requirement for *FIE* in the embryo and endosperm

Chapter 4. The requirement for *FIE* in the embryo and endosperm

4.1 Introduction

Analysis of *FIE* expression (by *in situ* hybridisation) shows that it is expressed in both embryo and endosperm during early seed development, but not in the developing seed coat (Spillane *et al.* 2000). At the octant stage of embryo development, *FIE* mRNA is localized in the embryo proper but not the suspensor. In the developing endosperm, *FIE* mRNA is weakly expressed in free endosperm nuclei covering the micropylar region of the embryo sac, but is abundant in the chalazal nodules. The distribution of *FIE* mRNA in developing seed is therefore consistent with it functioning in both early embryo and endosperm development after fertilization. Although aborting *fie* seed show defects in both endosperm and embryo development (sections 1.4 and 1.5), the embryo relies on the endosperm for nourishment and signals (section 1.3.1). This raises the possibility that embryo arrest could be an indirect consequence of defects in endosperm development.

Occasionally, *mea/mea* or *fis2/fis2* plants are obtained, but never *fie* homozygotes. Unlike *MEA* and *FIS2*, *FIE* is a single copy gene, so *mea* and *fis2* may just have a weaker phenotype due to redundancy, for example with *CLK* or *EMF2* respectively. Alternatively, *FIE* may have additional functions in the embryo, consistent with its expression there. In order to establish the requirements for *FIE* in the embryo and endosperm I used a strategy to target *FIE* expression to specific domains of the seed. For example, one aim was to express *FIE+* in the endosperm of a *fie*⁻ seed. This was intended to reveal the requirement for *FIE* in the embryo when the surrounding endosperm has *FIE+* activity.

4.2 Targeted gene expression strategy

Based on a method widely used in *Drosophila* (Brand and Perrimon, 1993), Jim Haseloff and colleagues developed a strategy for targeted gene expression in *Arabidopsis thaliana*. A collection of transgenic *Arabidopsis* lines was generated that stably express a chimeric transcriptional activator comprising of the DNA binding domain of the yeast activator GAL4 and a highly acidic portion of the herpes simplex virus protein VP16 (Sadowski *et al.* 1988). The GAL4-VP16 activator is used to coordinate expression of a linked *GREEN FLUORESCENT PROTEIN* (*GFP*) gene in particular cell types, and can in turn be used to drive target gene expression in those cells. A transformation vector was designed so that expression of the transcriptional activator, GAL4-VP16, would be dependent upon the fortuitous proximity of an *Arabidopsis* enhancer element when randomly inserted into the *Arabidopsis* genome by *Agrobacterium tumefaciens*-mediated transformation. However, it was found that due its high AT content, GAL4-VP16 is poorly expressed in *Arabidopsis*; the codon usage was therefore altered so that it is expressed more efficiently in *Arabidopsis* (http://www.plantsci.cam.ac.uk/Haseloff/GAL4/mGAL4.html). The *GFP* reporter gene is under the control of the *GAL4* upstream activating sequence (UAS) to give the 'driver' (figure 4.1a). However, it was found that due to a cryptic splice site *GFP* is not expressed well in *Arabidopsis* and that the small protein is dispersed throughout the cytoplasm and nucleoplasm. The sequence was therefore altered to produce the variant mGFP5 that expresses more efficiently in *Arabidopsis,* and is targeted to the endoplasmic reticulum (ER) to limit free-radical damage to the plant (Haseloff *et al.* 1997). GFP (originating from the jellyfish, *Aequorea victoria*) is ideal for use as a biological marker in living cells as it is small, relatively non-toxic and causes limited perturbation of the cells it is expressed in. When expressed in *Arabidopsis* the protein can be visualised directly by fluorescence and confocal microscopy. A chosen target gene can be placed under the control of GAL4 upstream activation sequences (UAS), introduced into plants by *Agrobacterium* mediated

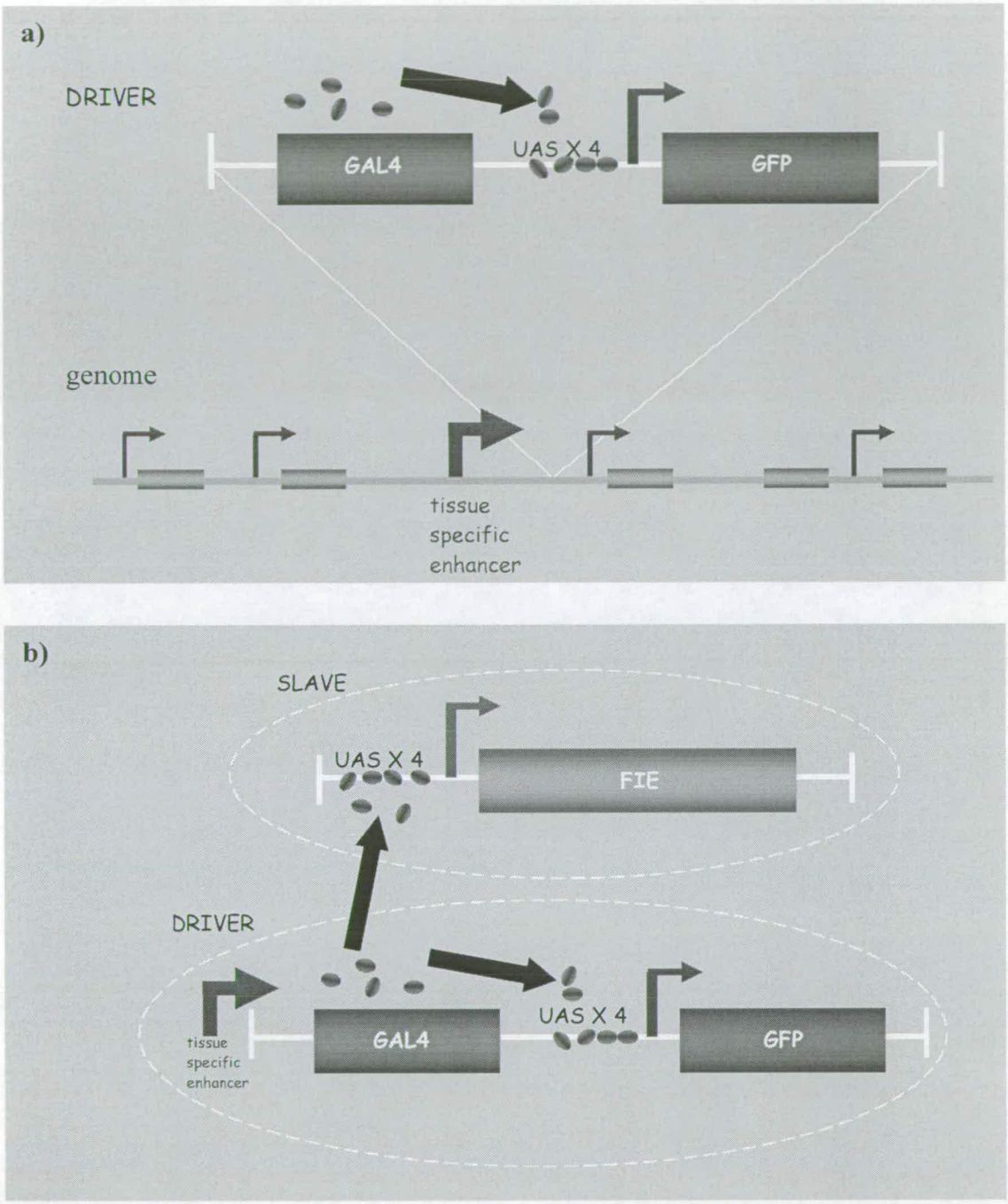


Figure 4.1 Targeted gene expression strategy. a) mGFP5 expression is dependent upon GAL4-VP16. When randomly inserted into the *Arabidopsis* genome GAL4-VP16 is driven by the fortuitous proximity of an *Arabidopsis* enhancer element, allowing for tissue specific GAL4-VP16 and mGFP5 expression. b) The driver line can be crossed into a slave line carrying the gene of interest (in this case *FIE*) under the control of GAL4 to allow for coordinated tissue specific expression of the gene of interest.

transformation, to give a 'slave', which is usually transcriptionally dependent on GAL4 activity (figure 4.1b).

4.3 Enhancer-trap lines conferring localised *GFP* expression during seed development

The collection of enhancer-trap lines generated by Jim Haseloff were screened by Fred Berger (ENS, Lyon) and Gwyneth Ingram (ICMB, Edinburgh) to identify individuals with distinct and stable patterns of GAL4-VP16 and GFP expression in the endosperm or embryo of the developing seed. I chose three driver lines for use, based on their localised GFP expression within specific regions of the developing seed (figures 4.2-4.4).

i. Embryo specific expression

Figure 4.2 shows the GAL4-VP16 enhancer trap line 9104. Expression of the ER-targeted variant mGFP5 is seen early on in seed development in the suspensor, and in some of the surrounding maternal tissue (figure 4.2a). Later during development, GFP expression is seen in the suspensor, in some of the surrounding tissue and in the embryo proper (figure 4.2b). GFP expression persists in the vasculature of the mature embryo (figure 4.2c). No expression was observed in the endosperm at any stage during development.

ii. Early endosperm expression

The GAL4-VP16 enhancer trap line K22 expresses the ER-targeted variant mGFP5 in the endosperm of the developing seed (figure 4.3). This highlights the ER-rich MCE and CZE and the NCD in the PE during early seed development (figure 4.3a-c). Nuclei appear as black circles surrounded by green nuclear envelopes. When the embryo reaches the heart and early torpedo stage in development, expression is restricted to the embryo-surrounding

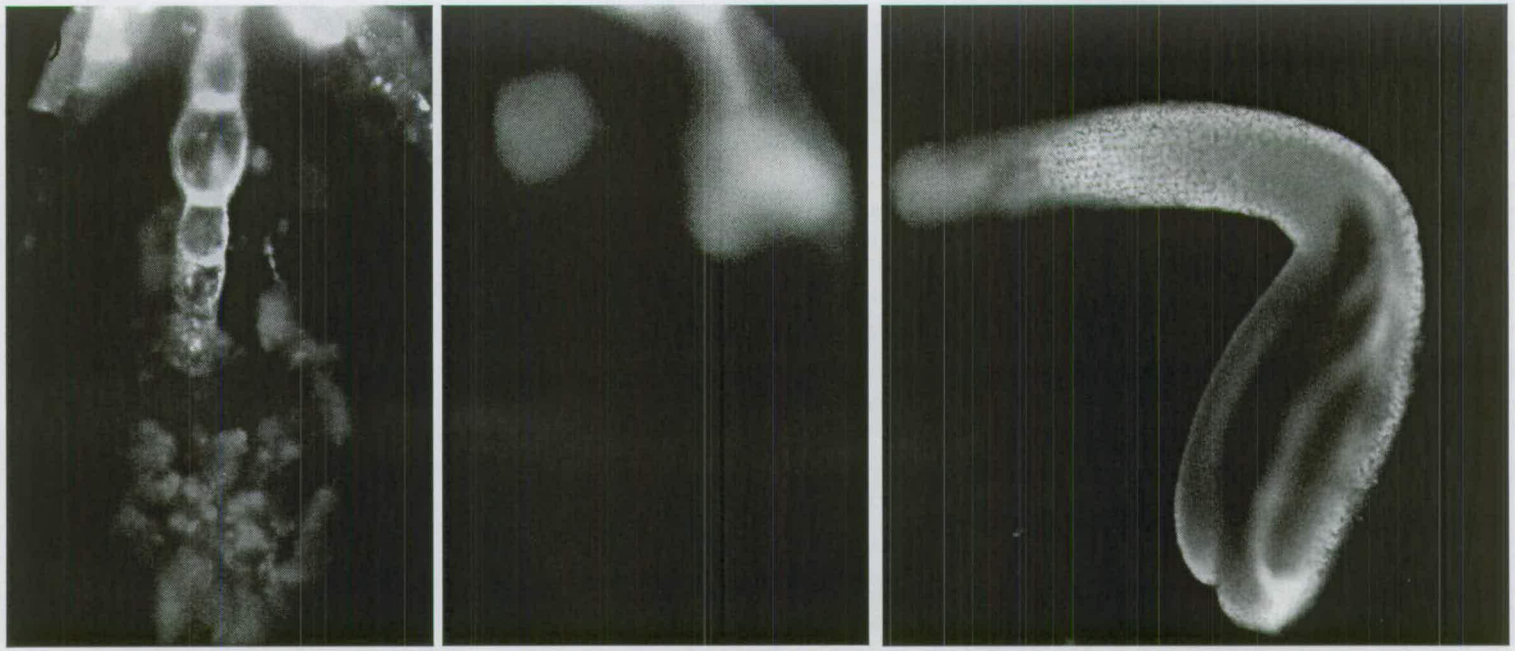


Figure 4.2 Embryo specific GFP expression in driver line 9104. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. a) Expression of the ER-targeted variant mGFP5 is seen early on in seed development in the suspensor, and in some of the surrounding maternal tissue. b) Later during development GFP expression is seen in the suspensor, in some of the surrounding tissue and throughout the embryo proper. c) The GFP expression persists in the vasculature of the mature embryo.

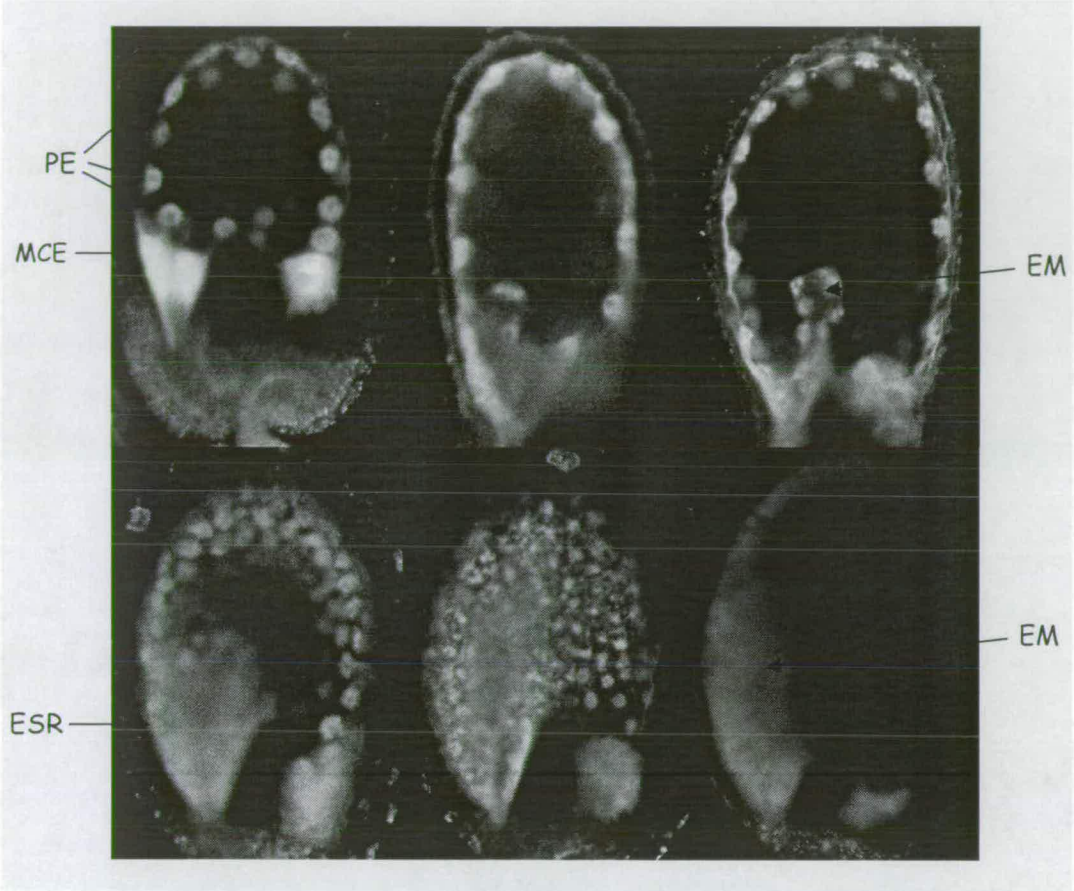


Figure 4.3 Early endosperm specific GFP expression in driver line k22. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. Expression of the ER-targeted variant mGFP5 is restricted to the endosperm of the developing seed. a)-c) The GFP highlights the ER-rich MCE, CZE and the NCD in the PE during early seed development. Nuclei appear as black circles surrounded by green nuclear envelopes. d) and e) GFP expression is restricted to the embryo-surrounding region (ESR) as the embryo reaches the heart and early torpedo stage in development. f) No expression is observed after the torpedo stage. MCE = mycropylar endosperm, CZE = chalazal endosperm, NCD = nuclear cytoplasmic domain, PE = peripheral endosperm, EM = embryo.

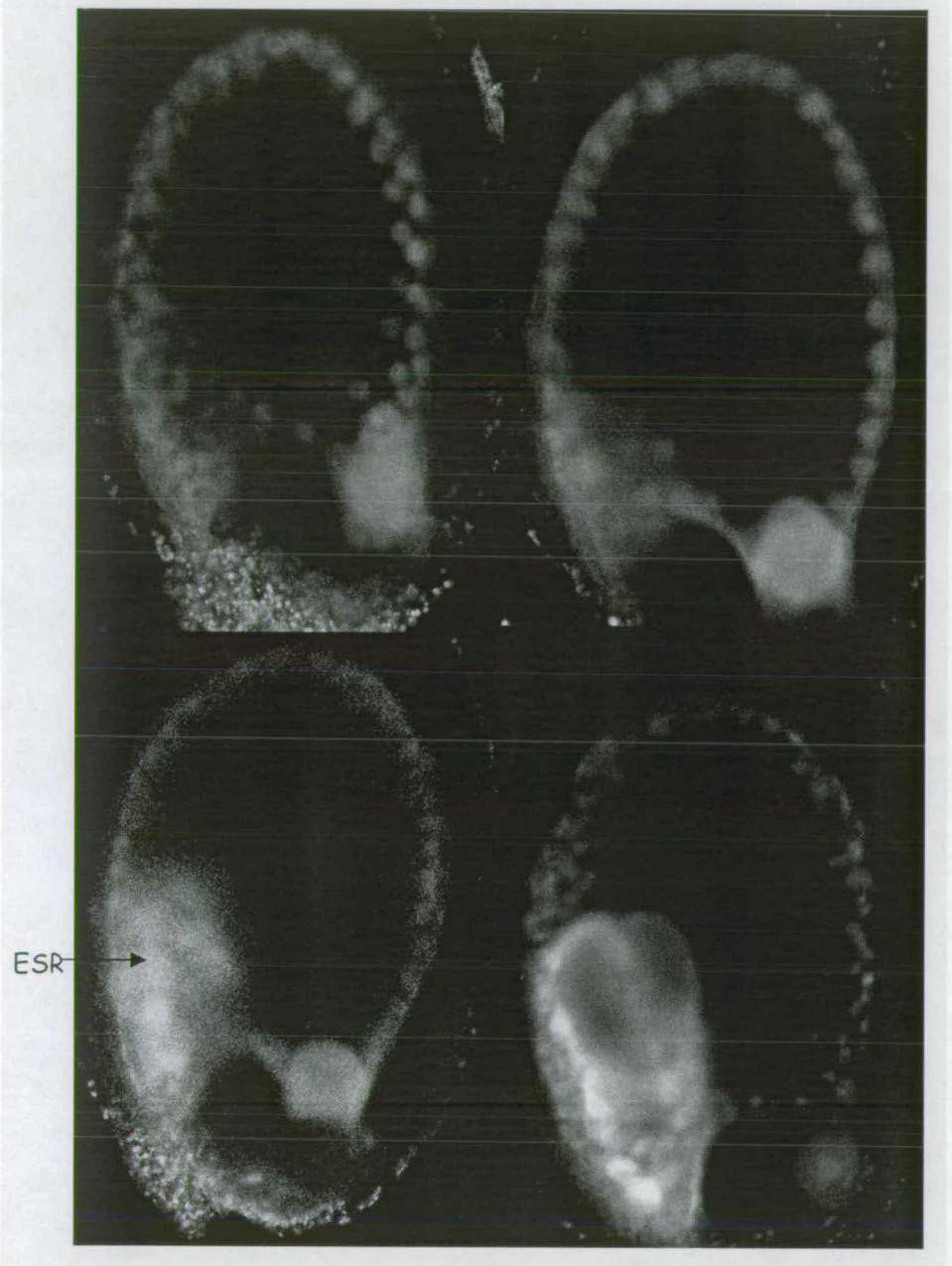


Figure 4.4 Expression of GFP in driver line 9185 during late endosperm development. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. a)-d) Show progressively later developmental stages. GFP expression is restricted to endosperm of the ESR. Expression begins after the heart stage and persists in the mature seed until desiccation. ESR = embryo surrounding region.

region (ESR) (figure 4.3d and e). No expression is observed after the torpedo stage (figure 4.3f).

iii. Late endosperm expression

The GAL4-VP16 enhancer trap line 9185 expresses the ER-targeted variant mGFP5 in the endosperm of the ESR of the developing seed (figure 4.4). Expression begins after the heart stage and persists in the mature seed until desiccation.

Seed development in these three driver lines appeared normal, confirming that the T-DNA insertion and GFP expression were without obvious phenotypic consequence (data not shown).

4.4 Expression of paternally inherited enhancer-trap lines

A delayed activation of the paternal genome during seed development has been reported by Vielle-Calzada *et al.* (2000). Using reciprocal crosses between wild type plants and enhancer trap lines carrying a GUS reporter gene they demonstrated an absence of paternal expression in the 20 loci tested up to 80 hours after pollination. As the loci were distributed throughout the genome and were functionally diverse they suggested that there is a genome-wide delayed activation of the paternal genome and so early embryo and endosperm development are mainly under maternal control. Weijers *et al.* (2001), however, show that for one transgene at least, paternal expression in embryos occurs as early as the two-cell stage. They measured uniparental gene activity by monitoring transgene expression in embryos from reciprocal crosses. They used a GUS reporter transgene driven by the promoter of the gene encoding the small-ribosomal-subunit protein 5A, which is strongly expressed during early seed development. Fertilization of wild type egg cells with

transgenic pollen resulted in detectable GUS expression in most embryos as early as the two-cell stage. They also looked at the early acting double-mutant *knolle keule*, in which cytokinesis is blocked from fertilization onwards, resulting in morphologically distinct mutant embryos as early as the one cell stage. They found that selfing *knolle keule* heterozygous plants gave the expected one-quarter of mutant embryos (*knolle keule* are closely linked) (Waizenegger *et al.* 2000), whereas almost all embryos from wild type pollination were normal. Weijers *et al.* (2001) concluded that functional paternal alleles are sufficient to rescue defective maternal alleles from as early as the two-cell stage. More recently, Baroux *et al.* (2001) used the more sensitive (compared with GUS) BARNASE reporter gene and found that several paternally inherited transgenes are expressed during very early embryogenesis (two-cell stage) but at low levels and are gradually activated, reaching the equivalent level of maternal expression around the mid-globular stage. They concluded that this phenomenon likely corresponds to a global paternal effect. Taken together these results suggest that contributions by the maternal and paternal genome to early seed development are not equivalent; there is a global mechanism for parental genome silencing/attenuation that is alleviated around the mid-globular stage of embryo development, but specific genes that are important in early development may be specifically activated earlier in development.

Yadegari *et al.* (2000) report that after pollinating wild type plants with *FIE::GFP* pollen, no GFP fluorescence is detected until the 100 NCD stage of endosperm development. Assuming that the *FIE::GFP* reporter faithfully mimics the expression of the endogenous *FIE* gene, this suggests that expression of the paternal *FIE* allele is delayed during seed development (a phenomenon which may be common to many genes, as discussed above). Paternal *FIE*+ alleles cannot rescue the seed lethality of maternal *fie*- alleles, suggesting that *FIE*+ activity is required very early in seed development, before the 100 NCD stage. This was clearly relevant for the rescue strategy used in this study, as it implied that *GAL4-VP16*

inherited paternally might not be active early enough in seed development. I therefore characterised the expression of the driver lines when paternally inherited.

If the driver lines used show late paternal expression of the transgene, then the expected frequency of rescue in seed of self pollinated hemizygotes will be reduced, as only maternal copies will be activated sufficiently early. The paternal activation of the GAL4 gene in the driver lines k22, 9185, and 9104 was investigated by emasculating wild type plants and fertilizing with pollen from the driver lines. GFP expression was not observed in the embryo specific line 9104 until three DAP (days after pollination) (globular-stage embryo) and thereafter was maintained throughout embryogenesis (figure 4.5). This indicates that the 9104 GAL4 transgene could be affected by the global paternal silencing/attenuation reported by Baroux *et al.* (2001) but is not paternally silenced beyond this early stage of development. It is possible that the transgene may be expressed before 3 DAP but at levels undetected by confocal microscopy. It is not clear if maternal expression begins before 3 DAP because the strong expression in surrounding maternal tissue masks the internal tissues. Expression of the K22 GFP marker gene was seen at 2 DAP (pre-globular stage embryo) (data not shown), and was expressed in a similar temporal and spatial pattern to homozygous selfed seed although at a lower level and disappeared at 5 DAP (figure 4.6), indicating that the k22 *GAL4* transgene may not be paternally silenced at all. Interestingly, the late endosperm marker 9185 showed no paternal *GFP* expression at any stage of seed development (figure 4.7). This suggested that the enhancer that drives expression of the *GFP* in line 9185 is imprinted such that only the maternal allele is expressed. As rather few imprinted genes have been identified in *Arabidopsis thaliana* it would be very interesting to isolate the DNA flanking the T-DNA insert in this line.

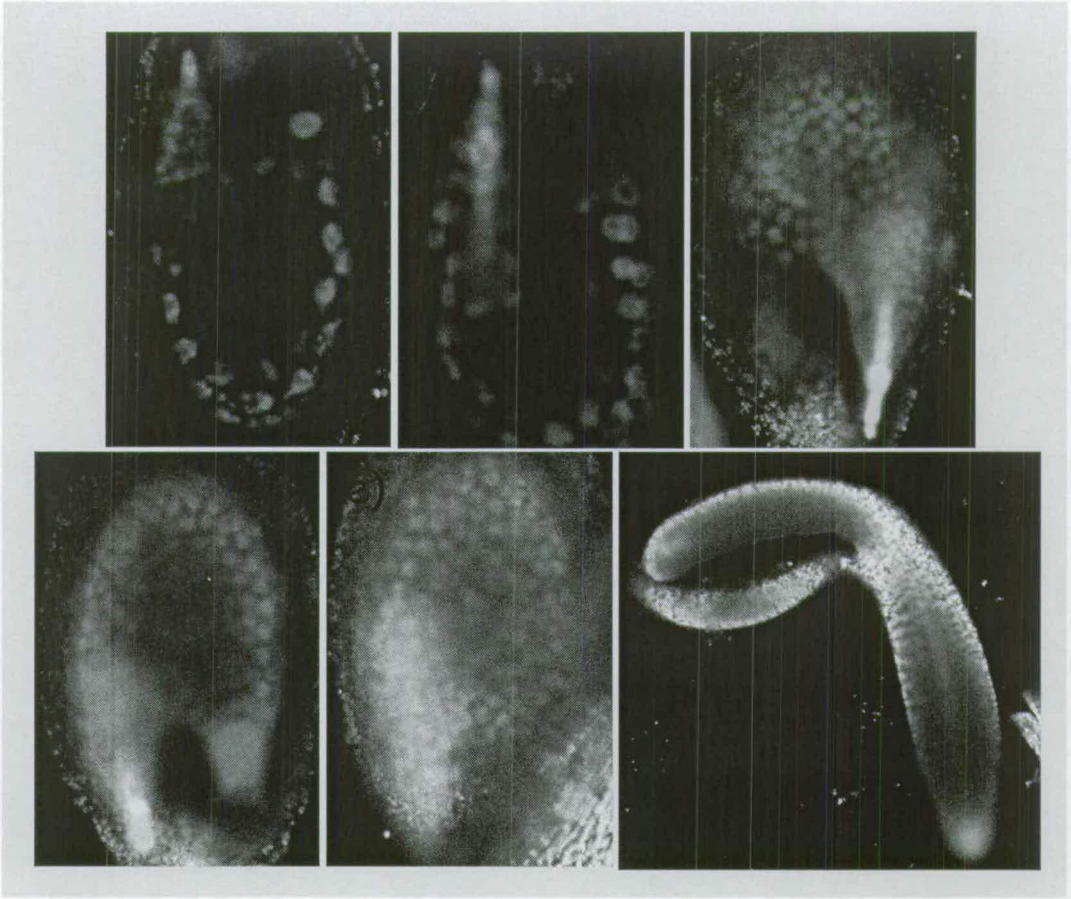


Figure 4.5 GFP expression from the paternal allele in driver line 9104. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. GFP fluorescence is observed as early as three DAP (days after pollination) (globular-stage) and was maintained throughout embryogenesis. a) 3 DAP, b) 4 DAP, c) 5 DAP, d) 6 DAP, e) 7 DAP, f) 8 DAP.

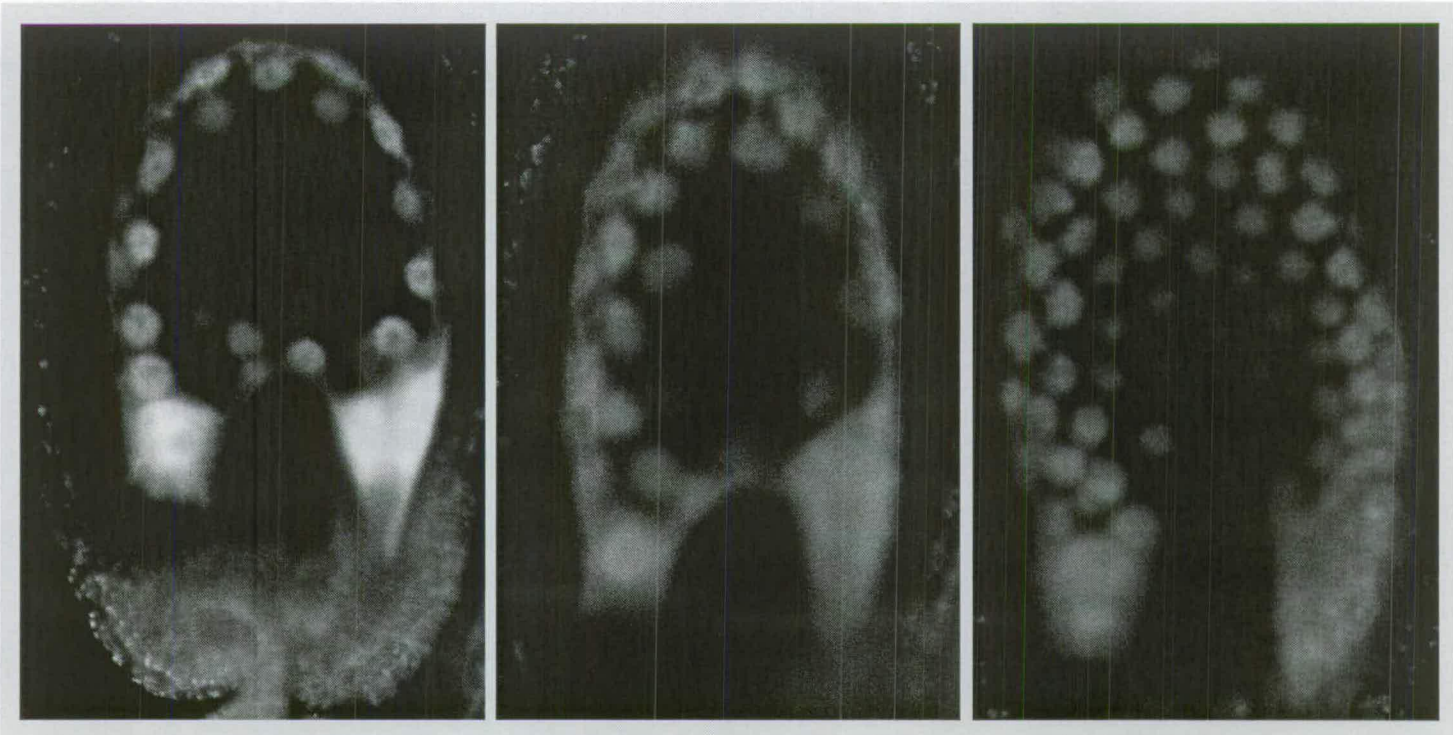


Figure 4.6 GFP expression from the paternal allele in driver line k22. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. Paternal GFP is expressed in a similar temporal and spatial pattern to homozygous selfed seed although at a lower level and disappeared at 5 DAP. a) Homozygous maternal and paternal expression 3 DAP. b) Paternal expression 3 DAP. c) Paternal expression 4 DAP.

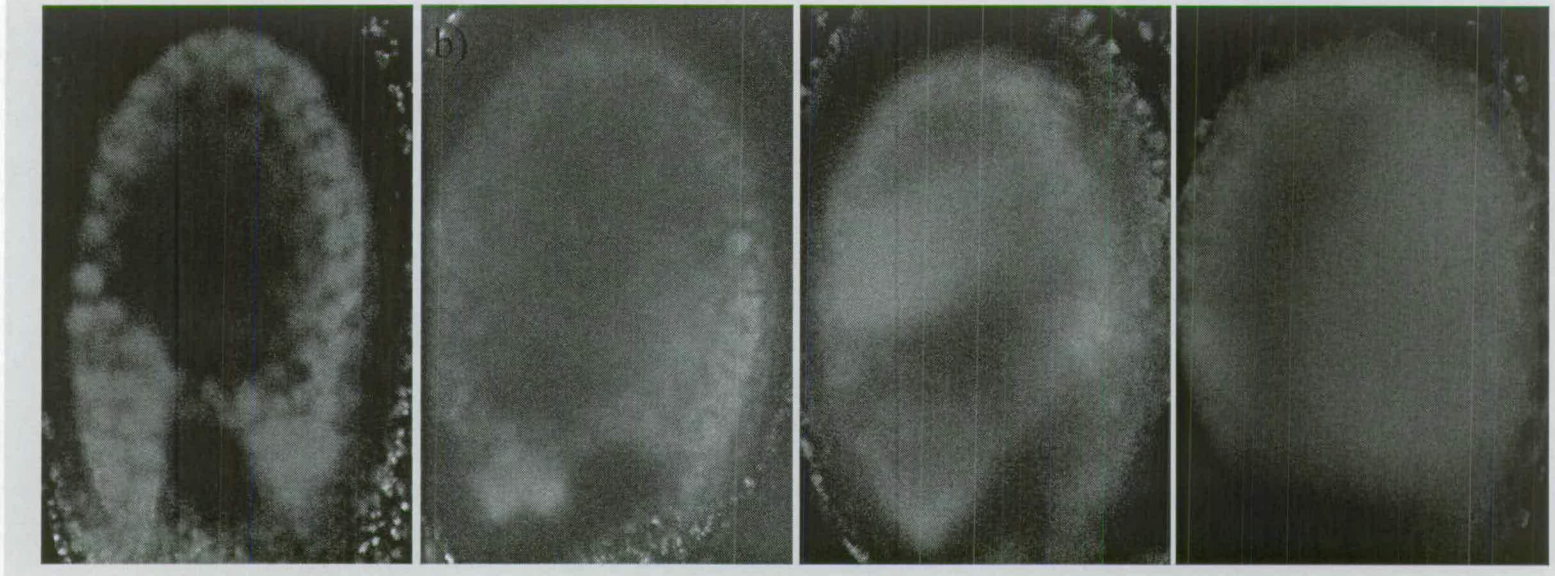


Figure 4.7 GFP expression from the paternal allele in driver line 9185. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. No paternal GFP expression was seen at any stage throughout seed development. a) 5 DAP (days after pollination). b) 6 DAP. c) 7 DAP. d) 8 DAP.

4.5 Effects of *fie* on driver line *GFP* expression

For some enhancer trap lines at least, *GFP* expression is perturbed in a *fie* mutant background. This might have consequences for the rescue strategy, particularly if the changes arise very early in seed development. Sørensen *et al.* (2001) reported ectopic *GFP* expression in driver line KS117 when expressed in *fie* seed. Heterozygous *fie*/+ plants, hemizygous for the KS117 T-DNA insertion were self-pollinated and the resulting seed analysed. In wild type seed GFP expression was restricted to the CZE at the globular and heart embryo stages. *fie* seed were characterized by a high and uniform expression of GFP marker in all parts of the endosperm. After the heart-stage of embryo development, GFP expression remained high in all parts of the endosperm, and the *fie* CZE cyst and nodules exhibited a very strong fluorescence relative to that of the wild type (Sørensen *et al.* 2001).

To investigate whether the driver lines k22, 9104, and 9185 showed similar changes, homozygous driver lines were fertilized with pollen from *fie*/+ heterozygotes. *fie*/+ F1 progeny were identified and the F2 seed examined for GFP expression. The embryo specific driver line (9104) showed no change in the pattern of GFP expression in the *fie* background. GFP is expressed in the aborting embryo throughout embryogenesis, and no expression is seen in the endosperm at any stage (figure 4.8). In a *fie* background, the early endosperm specific driver line (k22) is expressed throughout the CZE, MCE and PE in a similar pattern to wild type although at higher levels and more persistently, and no expression is seen in the embryo at any stage (figure 4.8). The endosperm-specific driver line (9185) confers GFP expression in the ESR of the aborting embryo in a *fie* background as well as the chalazal pole (figure 4.8). So it appears that the enhancer trapped in line 9185 shows ectopic expression in a *fie* background, although the expression is still endosperm specific, whereas the GFP expression patterns of driver lines 9104 and k22 are not altered spatially and are consistent with the morphological changes observed in *fie* aborting seed.

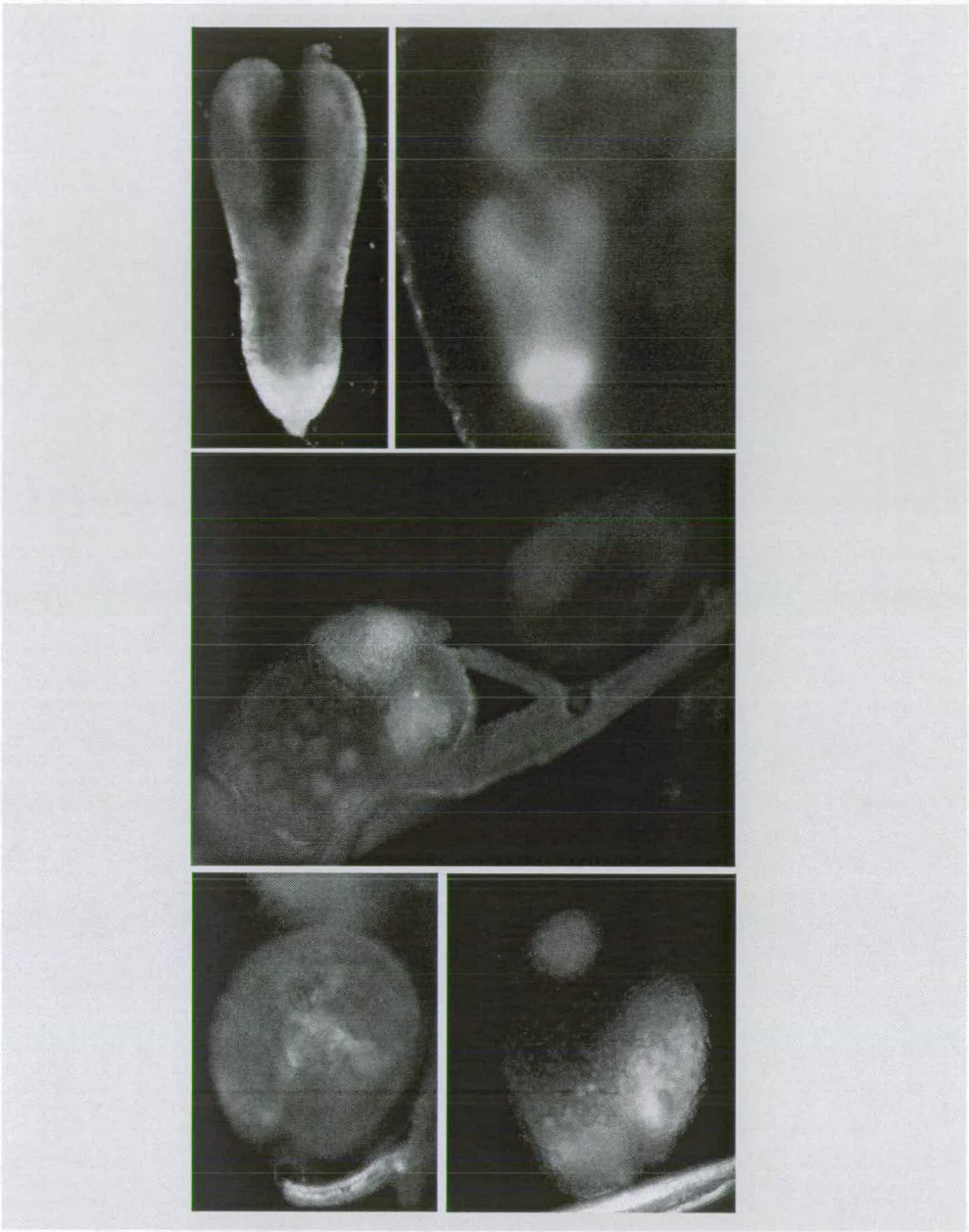


Figure 4.8 Effect of *fie* on driver line expression patterns. Confocal sections of developing seed. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. Seed from each line are from the same silique. a) to b) Embryo specific line 9104. a) *FIE*+ embryo. b) *fie*⁻ embryo. No ectopic expression of the GFP marker is observed. c) Early endosperm specific line k22. Right hand seed (*FIE*+) shows an early walking stick stage embryo and no GFP expression in the endosperm. Left hand seed (*fie*⁻) shows endosperm at a earlier developmental stage than wild type sibling, with strong GFP expression in all areas of the endosperm. d) to e) Late endosperm specific line 9185. d) *FIE*+ seed, GFP is expressed in the ESR. e) *fie*⁻ seed. GFP is expressed in the ESR of the aborting embryo in a *fie* background as well as the chalazal pole.

4.6 Generation of *UAS::FIE* 'slave' lines

The *FIE* cDNA was subcloned between a *GAL4 UAS* promoter and a *NOS* (nopaline synthase) terminator. This construct was introduced into the pBIB-Hyg^R binary vector (Becker 1990) conferring Hygromycin resistance as a plant selectable marker (which allows it to be followed independently of the kanamycin resistant driver), and introduced into *Agrobacterium*. The construct was then introduced into *fie-2/+* plants (*Ler* background) by *Agrobacterium* mediated transformation. The transformed T1 seed were then germinated in sterile tissue culture media containing Hygromycin to select primary transformants. Around 30 transformed T1 plants were recovered, some of which carried the *fie-2* mutation. T2 seed were then sown on plant tissue culture plates containing Hygromycin and segregation data recorded to estimate the number of transgene insertion sites. T2 Hyg^R plants carrying T-DNA insertions at one (or in one case, two) loci (table 4.1) were crossed as males to the endosperm specific driver lines k22 and 9185, and to the embryo specific driver line 9104.

ID	*FIE* genotype	Hyg sensitive	Hyg resistant	Ratio s:r
13.1	*fie*/+	4	9	1:2
14.2	*fie*/+	2	35	1:18
27.1	*fie*/+	10	19	1:2
14.1	+/+	16	35	1:2
14.3	+/+	24	50	1:2
14.5	+/+	15	38	1:2

Table 4.1 T2 segregation data. The ratios are consistent with single locus insertions, except for line 14.2, which likely has inserts at two unlinked loci.

Unfortunately, it was later established that slave line 13.1 did not contain the T-DNA insertion, and that slave line 27.1 conferred an embryo lethal phenotype (figure 4.9). Line 14.2 however, displayed 50% wild type and 50% aborting *fie* seed, and both phenotypes appeared unchanged by the T-DNA insertion (figure 4.9). All the data below corresponds to the slave line 14.2.

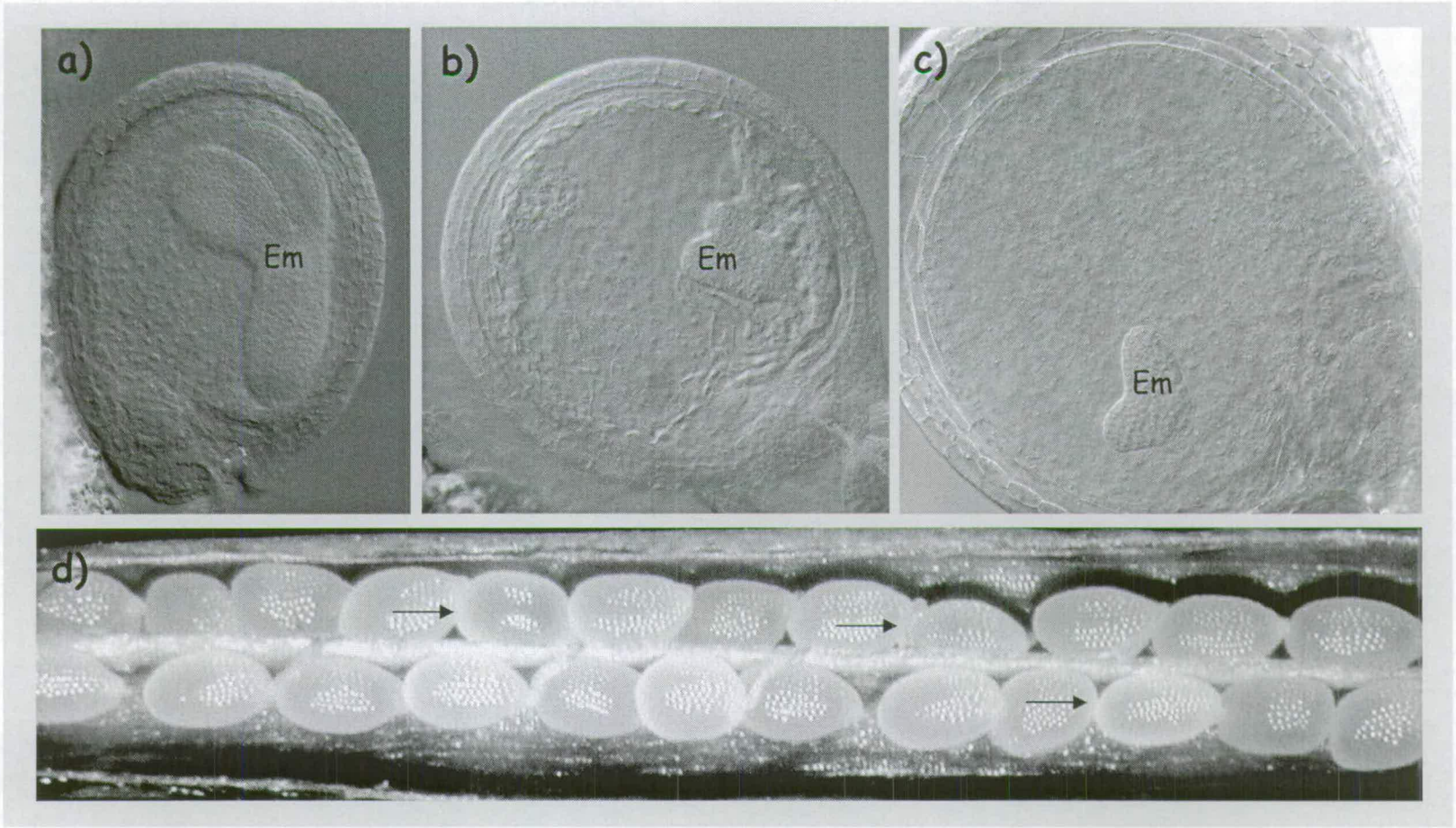


Figure 4.9 Seed development in *UAS::FIE* slave lines. a) to c) DIC light microscopy of cleared developing seed. a) Line 14.2, *FIE+* seed. b) Line 14.2, *fie-* seed. c) and d) Line 27.1, *FIE+* seed. Segregated 25% seed with an embryo lethal phenotype. Shown here is an abnormal heart shaped embryo. d) Line 27.1 dissected silique. Arrows indicate seed displaying embryo lethality and normal endosperm. Em = embryo

4.7 Experimental strategy

Slave lines (14.2) carrying the *UAS::FIE* insert and segregating for *fie-2* were crossed as males to the three driver lines described above. F1 plants were then selected for driver (GFP fluorescence), slave (a Hygromycin resistance ratio of 3:1), and the *fie* phenotype of seed abortion. Plants carrying all three traits were selfed, and the resulting F2 seed were again selected on Hygromycin to monitor the presence of the *UAS::FIE* insertion. F2 plants were also screened for GFP expression and the *fie* phenotype. Seed homozygous for both driver and slave insertions were identified and seed from a F3 family (F4 generation) carrying the embryo specific driver were examined for aberrant seed abortion ratios and developing seed phenotypes. However, due to time constraints homozygous F3 plants of the two endosperm specific driver lines were not examined, only the segregating F2 families (F3 seed) were characterised.

Targeted expression of the *FIE* transgene specifically in the embryo or endosperm in a *fie* background could give rescue of the seed abortion phenotype. In this case *fie/+* plants homozygous for both driver and slave transgenes will show 100% seed viability. If hemizygous for the driver transgene, then it is likely that only seed inheriting the transgene maternally will be rescued, giving 25% seed abortion, whilst plants hemizygous for both transgenes would display around 37.5% seed abortion. Other factors to consider are that the driver and slave transgenes may be linked to the *FIE+* allele on chromosome III, skewing the abortion ratios, or that the GAL4 UAS promoters may become silenced, especially in advanced generations of homozygous lines. This can occur due to post-transcriptional gene silencing (PTGS) of the transgenes, a phenomenon that can be induced by highly transcribed sense transgenes (S-PTGS) (Vaucheret *et al.* 1997 and Que *et al.* 1997) or by transgene loci producing double-stranded RNA (dsRNA) due to the presence of inverted repeats (IR-PTGS) (Stam *et al.* 1998, Waterhouse *et al.* 1998, Chuang and Meyerowitz 2000, and Sijen *et al.*

2001). It should also be considered that the F3 and F4 seed characterised in this study is of a mixed background (ecotypes C24 and L*er*), and as such may harbour genes modifying the *fie* phenotype (some such modifiers have been reported for the *mea* phenotype in different ecotypes (Vielle-Calzada *et al.* 1999)).

4.8 Seed abortion ratios

i. Embryo specific

49 F3 plants segregating for *fie*, homozygous for the driver line 9104 *GFP* transgene and hemizygous for *UAS::FIE* (14.2) were examined. 61% (30/49) displayed a seed abortion phenotype, suggesting they carried the *fie-2* mutation. This percentage is not significantly different from 50% ($\chi 2$ = 2.47, 1 d.f., P = < 0-5), signifying that the *fie*+ mutation is segregating with the expected ratio of 1:1. 17% (5/30) of the *fie*/+ plants displayed significantly lower abortion ratios compared with the expected 50% (table 4.2), some as low as 40%, indicating that some rescue of the *fie*- phenotype may be occurring.

ii. Early endosperm specific

48 F2 plants segregating for *fie*, hemizygous for both driver line k22 *GFP* transgene and *UAS::FIE* were examined. 60% (29/48) displayed a seed abortion phenotype, suggesting they carried the *fie-2* mutation. This percentage is not significantly different from 50% ($\chi 2$ = 2.08, 1 d.f., P < 0.1), again suggesting that the *fie*+ mutation is segregating in the expected ratio of 1:1. Of the *fie*-/+ plants, 21% (6/29) displayed significantly lower abortion ratios compared with the expected 50% (table 4.3), some as low as 29%, indicating that some rescue of the *fie*- phenotype may be occurring.

Overall there is a lower frequency of driver line *GFP* expression seen, which could be attributed to silencing of the transgene, but is most probably due to mis-scoring, as the

Family number	Aborted	Wild type	Total	Percent aborted	χ^2(3.d.p)	P<
4/1	182	205	387	47	1.367	
4/5	163	249	412	40	17.951	0.005
4/6	207	239	446	46	2.296	
4/7	180	186	366	49	0.098	
4/9	224	281	505	44	6.434	0.025
4/12	158	181	339	47	1.560	
4/13	122	138	260	47	0.985	
4/15	**188**	**193**	**381**	**49**	**0.066**	
4/16	118	158	276	43	5.797	0.025
4/18	171	204	375	46	2.904	
4/19	130	169	299	43	5.087	0.025
4/21	157	174	331	47	0.873	
4/22	**173**	**203**	**376**	**46**	**2.394**	
4/26	91	105	196	46	1.000	
4/29	101	93	194	52	0.330	
4/30	99	86	185	54	0.914	
4/31	115	130	245	47	0.918	
4/32	95	120	215	44	2.907	
4/33	100	101	201	50	0.005	
4/34	113	112	225	50	0.004	
4/35	73	68	141	52	0.177	
4/36	**114**	**130**	**244**	**47**	**1.049**	
4/38	104	114	218	48	0.459	
4/43	112	133	245	46	1.800	
4/44	111	127	238	47	1.076	
4/45	131	133	264	50	0.015	
4/46	122	152	274	45	3.285	
4/48	100	138	238	42	6.067	0.025
4/49	122	140	262	47	1.237	
4/50	143	140	283	51	0.032	

Table 4.2. Seed abortion ratios for F4 seed from crosses involving the embryo specific driver line (9104). χ^2 values were calculated for an expected abortion ratio of 50%. Probability (P) that observed results differ from expected is only shown for those with significant differences (P< 0.05). This F3 family was homozygous for both driver and slave transgenes. Plants segregating a *fie* phenotype were identified by examining developing seed. Those emboldened did not display abnormally large seed. (*FIE+* plants are not shown). The plants above are all siblings of progeny of a single F2 plant, i.e. 4/36 is plant 36 in family 4.

***fie*- plants**

Plant family	Driver transgene	Aborted	Wild type	Total	Percent aborted	χ^2 (3 d.p.)	P<
30/1	**hemizygous**	**70**	**66**	**136**	**51**	**0.118**	
30/4	hemizygous	117	123	240	49	0.150	
30/5	hemizygous	55	122	177	31	25.361	0.005
30/6	hemizygous	130	141	271	48	0.446	
30/7	hemizygous	250	288	538	46	2.684	
30/10	**hemizygous**	**105**	**98**	**203**	**52**	**0.241**	
30/12	hemizygous	98	124	222	44	3.045	
30/14	hemizygous	103	115	218	47	0.661	
30/15	hemizygous	131	174	305	43	6.062	0.025
30/16	**none**	**134**	**133**	**267**	**50**	**0.004**	
30/17	hemizygous	120	110	230	52	0.435	
30/18	hemizygous	149	140	289	52	0.280	
30/19	**hemizygous**	**118**	**106**	**224**	**53**	**0.643**	
30/20	none	246	274	520	47	1.508	
30/26	homozygous	66	162	228	29	40.421	0.005
30/27	hemizygous	80	87	167	48	0.293	
30/28	hemizygous	66	90	156	42	3.692	
30/31	homozygous	126	219	345	37	25.070	0.005
30/33	hemizygous	242	316	558	43	9.814	0.005
30/34	hemizygous	61	62	123	50	0.008	
30/36	hemizygous	117	113	230	51	0.070	
30/40	**none**	**114**	**116**	**230**	**50**	**0.017**	
30/41	hemizygous	105	102	207	51	0.043	
30/42	homozygous	60	147	207	29	36.565	0.005
30/44	hemizygous	73	84	157	47	0.771	
30/46	none	111	136	247	45	2.530	
30/48	hemizygous	91	101	192	47	0.521	
30/49	homozygous	124	127	251	49	0.036	
30/50	**none**	**155**	**158**	**313**	**50**	**0.029**	

***FIE*+ plants**

30/2	**none**	**0**
30/3	**none**	**0**
30/9	**none**	**0**
30/11	**hemizygous**	**0**
30/13	**hemizygous**	**0**
30/21	**none**	**0**
30/22	**hemizygous**	**0**
30/23	**none**	**0**
30/25	**none**	**0**
30/29	**hemizygous**	**0**
30/30	**hemizygous**	**0**
30/32	**none**	**0**
30/35	**hemizygous**	**0**
30/37	**none**	**0**
30/38	**none**	**0**
30/39	**none**	**0**
30/43	**homozygous**	**0**
30/45	**none**	**0**
30/47	**hemizygous**	**0**

Table 4.3. Seed abortion ratios for F3 seed from crosses involving the early endosperm specific driver line (k22). χ^2 values were calculated for an expected abortion ratio of 50%. Probability (P) that observed results differ from expected is only shown for those with significant differences (P> 0.05). Those emboldened did not display abnormally large seed.

developing seed were screened at a young age when they could easily be damaged during silique dissection.

iii. Late endosperm specific

48 F2 plants segregating for *fie*, hemizygous for both driver line 9185 *GFP* transgene and *UAS::FIE* were examined. 58% (28/48) displayed a seed abortion phenotype, indicating they carried the *fie-2* mutation. Again, this percentage is not significantly different from 50% ($\chi 2$ = 1.33, 1 d.f., P < 0.5), signifying that the *fie+* mutation is segregating in the expected ratio of 1:1. Of the *fie-/+* plants, 22% (6/27) displayed significantly lower abortion ratios compared to the expected 50% (table 4.4), some as low as 38%, suggesting that some rescue of the *fie-* phenotype may be occurring.

2/27 (7.4%) of the plants displaying a *fie* phenotype appeared homozygous for the driver transgene, compared with 10/20 (50%) of the *FIE+* plants (table 4.4). Although this is not significantly different from the expected Mendelian 1:2:1 segregation of the driver transgene in the *fie-/+* plants ($\chi 2$ = 3.111, 2 d.f. P< 0.5), it is significantly different in the *FIE+* set ($\chi 2$ = 8.8, 2d.f., P< 0.005), suggesting that the driver line 9185 transgene might be linked to the *FIE+* allele on chromosome III.

If the *fie-* phenotype of embryo abortion was due to the absolute requirement of *FIE* in either the embryo or endosperm, one might expect to see a difference in the proportion of plants displaying lower abortion ratios depending on each driver line used. For example, if *FIE* is not required in the endosperm, but only in the embryo, one would expect no rescue of embryo abortion in the crosses involving the endosperm specific drivers, k22 and 9185. However, the fact that *fie* rescue was observed with all three driver lines suggests that either *FIE* is required in at least one of the tissues, or that the *UAS::FIE* transgene in the slave line 14.2 is inserted in a context that allows for some basal expression regardless of GAL4-VP16

***fie*⁻ plants**

Plant family	Driver transgene	Aborted	Wild type	Total	Percent aborted	χ^2 (3d.p.)	P<
230/1	hemizygous	129	181	310	42	8.723	0.005
230/2	**none**	**127**	**144**	**271**	**47**	**1.066**	
230/4	**hemizygous**	**86**	**81**	**167**	**52**	**0.150**	
230/5	hemizygous	154	200	354	44	5.977	0.0.25
230/6	hemizygous	197	210	407	48	0.415	
230/8	**none**	**127**	**151**	**278**	**46**	**2.072**	
230/10	hemizygous	149	237	386	39	20.062	0.005
230/15	**hemizygous**	**165**	**169**	**334**	**49**	**0.048**	
230/16	none	93	108	201	46	1.119	
230/18	**hemizygous**	**178**	**213**	**391**	**46**	**3.133**	
230/19	hemizygous	108	156	264	41	8.727	0.005
230/21	**hemizygous**	**191**	**174**	**365**	**52**	**0.792**	
230/22	hemizygous	115	114	229	50	0.004	
230/23	**none**	**104**	**84**	**188**	**55**	**2.128**	
230/24	**none**	**96**	**90**	**186**	**52**	**0.194**	
230/25	hemizygous	112	135	247	45	2.142	
230/26	**homozygous**	**146**	**140**	**286**	**51**	**0.126**	
230/27	hemizygous	199	202	401	50	0.022	
230/28	hemizygous	135	128	263	51	0.186	
230/30	hemizygous	142	200	342	42	9.836	0.005
230/31	**none**	**185**	**179**	**364**	**51**	**0.099**	
230/37	homozygous	121	126	247	50	0.101	
230/39	**hemizygous**	**103**	**104**	**207**	**50**	**0.005**	
230/40	hemizygous	118	193	311	38	18.087	0.005
230/42	hemizygous	176	192	368	48	0.696	
230/47	none	138	143	281	49	0.089	
230/49	**hemizygous**	**115**	**117**	**232**	**50**	**0.017**	

***FIE*⁺ plants**

230/3	**homozygous**	**0**
230/7	**homozygous**	**0**
230/11	**none**	**0**
230/12	**homozygous**	**0**
230/13	**none**	**0**
230/14	**homozygous**	**0**
230/17	**none**	**0**
230/20	**homozygous**	**0**
230/32	**hemizygous**	**0**
230/33	**none**	**0**
230/34	**hemizygous**	**0**
230/35	**hemizygous**	**0**
230/38	**hemizygous**	**0**
230/41	**none**	**0**
230/43	**none**	**0**
230/44	**homozygous**	**0**
230/45	**homozygous**	**0**
230/46	**homozygous**	**0**
230/48	**homozygous**	**0**
230/50	**homozygous**	**0**

Table 4.4. Seed abortion ratios for F3 seed from crosses involving the late endosperm specific driver line (9185). χ^2 values were calculated for an expected abortion ratio of 50%. Probability (P) that observed results differ from expected is only shown for those with significant differences (P> 0.05). Those emboldened did not display abnormally large seed.

activation. However, as some seed abortion percentages are lowest (29%) in the early endosperm specific cross, it suggests that *FIE* might be required in the early endosperm in particular in wild type plants. However, further crosses with independent *UAS::FIE* slave lines would need to be carried out to support this.

4.9 Novel phenotype in lines carrying *UAS::FIE*

In all three crosses, most of the plants (90% for 9104, 79% for k22, 56% for 9185, tables 4.2 to 4.4) segregating for *fie* produced some seed which were larger than wild type, and were often misshapen (figures 4.10 to 4.12). These large seed often appeared to remain green longer than their *FIE+* siblings in the same silique, which were undergoing desiccation. A high proportion of these plants also produced seed that appeared to have aborted, (they were brown in colour and had collapsed) but contained a much larger embryo than aborted seed observed in *fie-2* seed (figures 4.10 to 4.12). Abnormally large seed were never observed in the *FIE+* proportion. Also, no abnormal seed were observed in F3 seed from crosses lacking the *fie-2* mutation (data not shown). This suggests that this novel phenotype may be due to a partial rescue of *fie*$^{-}$ seed, and not due to a dosage effect as a result of expressing an extra copy of *FIE*.

The developing seed were cleared and examined using DIC (Nomarski Differential Interference Contrast) light microscopy to examine the embryo and endosperm. In all three crosses approximately half the seed appeared wild type, the rest of the seed showed some abnormalities ranging from typically *fie*-like to seed containing late aborting embryos but with developmentally delayed endosperm (figures 4.13 to 4.15). In the cross segregating for the k22 driver line insertion, aborting embryos with cellularized PE were also sometimes observed (figure 4.14d). In the cross segregating for the late endosperm specific 9185 driver insertion, seed containing embryos that have aborted at a later stage than *fie-2* embryos

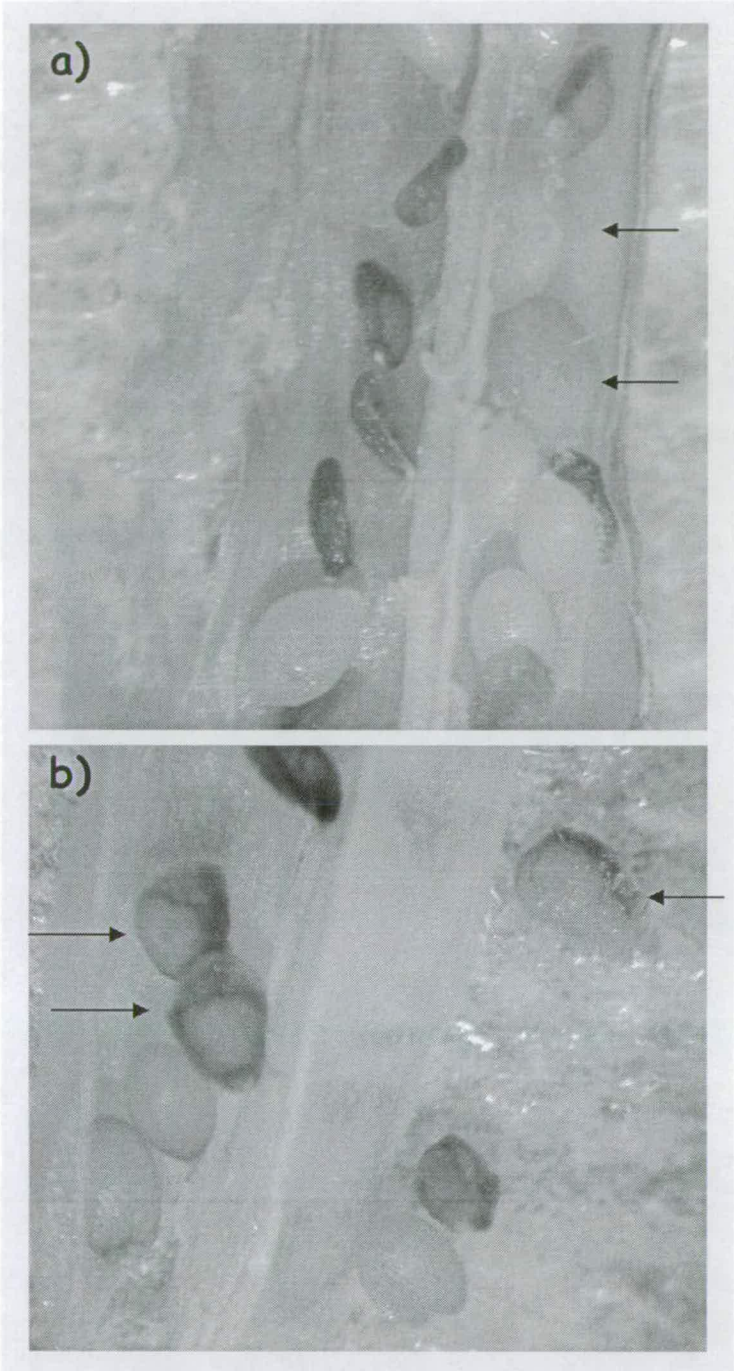


Figure 4.10 Effects of targeted expression of *FIE* in the embryo. Dissected siliques from F3 family 4 described in table 4.2. a) Plant 4/16. Arrows indicate large green developing seed. b) Plant 4/18. Arrows indicate aborting seed containing *fie* embryos that are larger and at later developmental stages than is normal for *fie*.

Figure 4.11 Effects of targeted expression of *FIE* in the early endosperm. Dissected siliques showing developing seed from F2 family 30 described in table 4.3. a) Plant 30/44. b) Plant 30/17. Black arrows indicate large misshapen seed, red arrows indicate aborting seed containing *fie* embryos that are larger and at later developmental stages than is normal for *fie*.

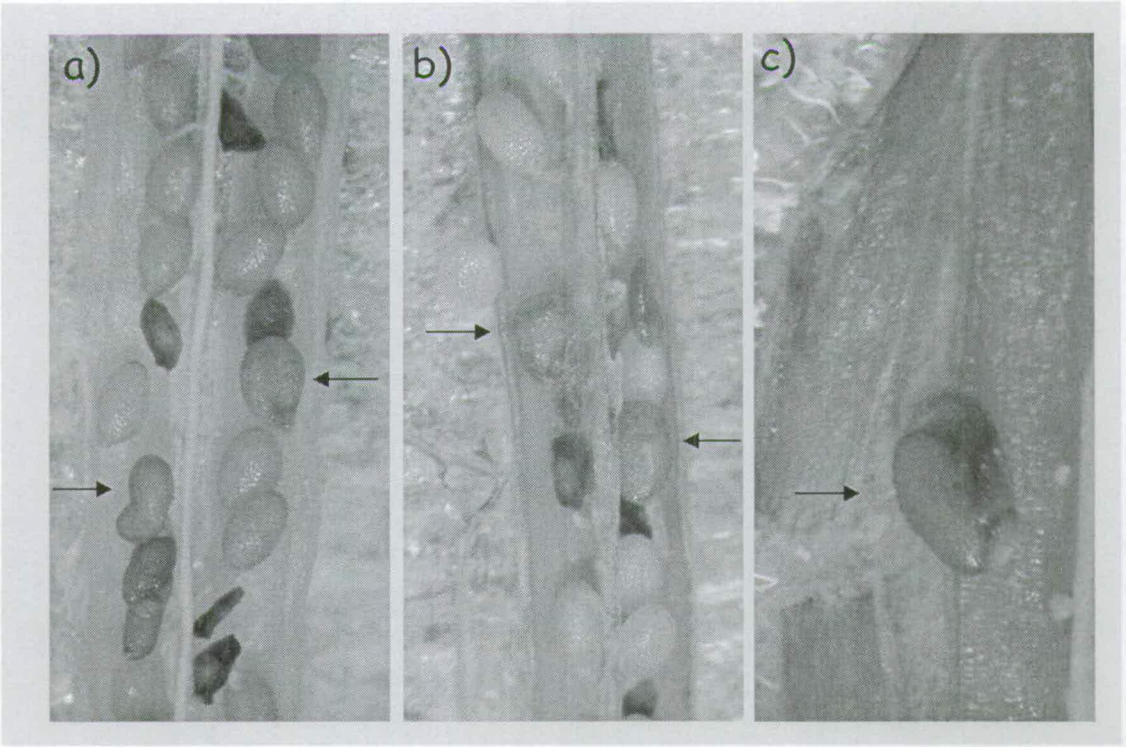


Figure 4.12 Effects of targeted expression of *FIE* in the late endosperm. Dissected siliques showing developing seed from F2 family 230 described in table 4.4. a) Plant 230/10. Arrows indicate large seed. b) Plant 230/16. Arrows indicate large seed. c) Line 230/30. Arrow indicates aborting seed containing *fie* embryos that are larger and at later developmental stages than is normal for *fie*.

.

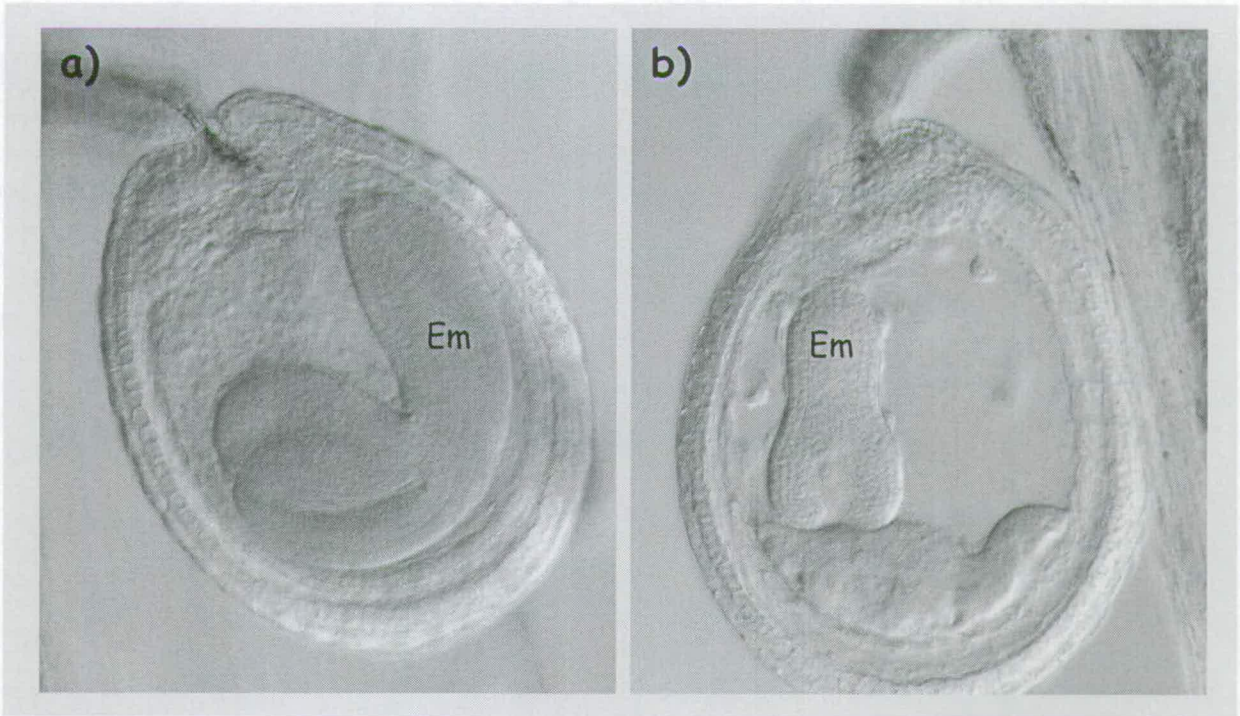


Figure 4.13 Effects of targeted *FIE* expression in the embryo. Developing seed from F3 family 4 described in table 4.2. DIC Light microscopy of cleared developing seed (Plant 4/44). Seed are from the same silique. a) *FIE*+ seed. b) *fie*⁻ seed. The endosperm is typical of *fie* mutants but the embryo is developmentally more advanced than is usually seen in aborting *fie* seed and has reached the torpedo stage of development. Em = embryo.

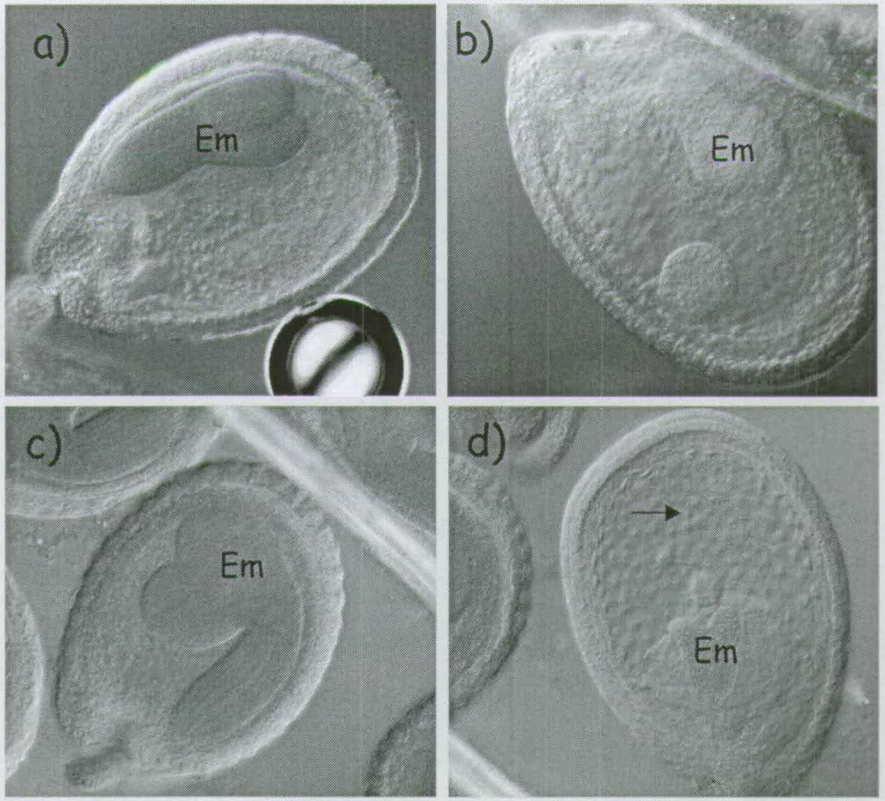


Figure 4.14 Effects of targeted *FIE* expression during early seed development. Developing seed from F2 family 30 described in table 4.3. DIC Light microscopy of cleared developing seed. a) and b) Plant 30/28, c) and d) Plant 30/34. Seed of each line are from the same silique. a) and c) *FIE*+ seed. b) and d) *fie*⁻ seed. b) The embryo and endosperm appear developmentally more advanced than is usually seen in aborting *fie* seed, but contains an ectopic chalazal cyst typical of *fie* mutants. d) Arrow indicates partially cellularised endosperm. Em = embryo.

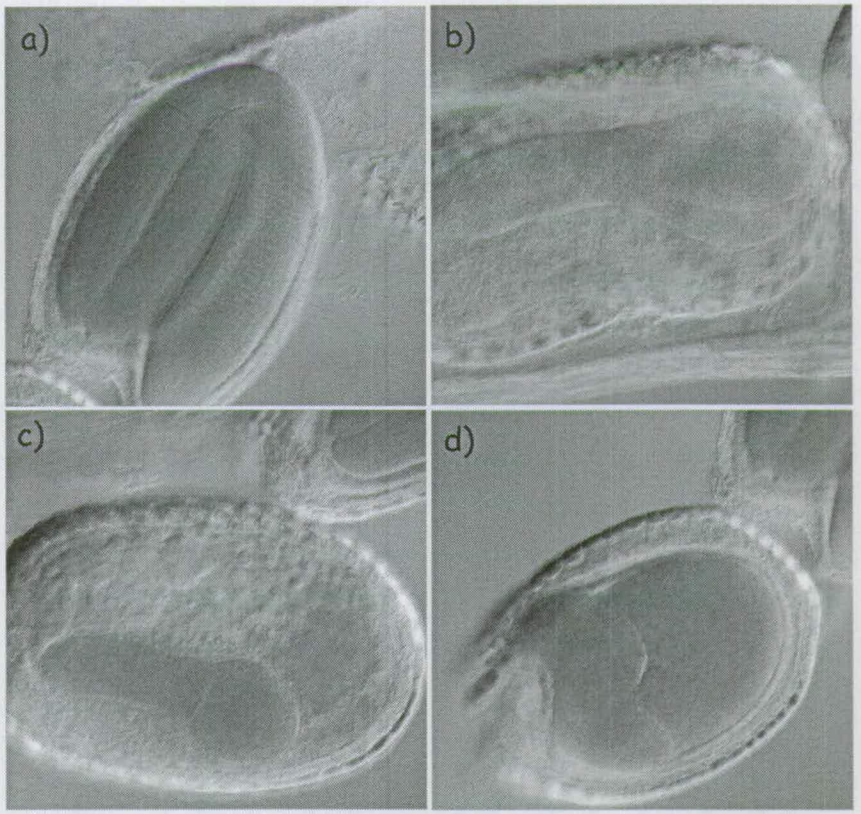


Figure 4.15 Effects of targeted *FIE* expression during late endosperm development. Developing seed from family 230 described in table 4.4. DIC light microscopy of cleared developing seed. (Plant 230/25.) Seed are from the same silique. a) *FIE*+ seed with mature embryo. b-d) *fie*⁻ seed, the embryo and endosperm are developmentally more advanced than is usually seen in aborting *fie* seed. d) Shows misshapen embryo. Em = embryo.

(figures 4.15b and c) and occasionally appearing morphologically abnormal (figure 4.15d) were observed. The endosperm of the late aborting seed also appears more developmentally advanced than in *fie*-2 seed, although often contained a large chalazal cyst typical of *fie*.

4.10 Rescue of *fie/fie* homozygotes by targeted *FIE* expression

In all three lines there are plants displaying a significantly reduced seed abortion ratio, suggesting that some *fie* embryos may be escaping seed abortion. F3 or F4 seed from the three crosses were therefore sown onto plant tissue culture plates with or without Hygromycin. Morphologically abnormal plants were genotyped to identify *fie* homozygotes by using CAPS (cleaved amplified polymorphic sequence) markers and sequencing analysis. A small proportion from each of the three crosses produced developmentally abnormal seedlings some of which were later confirmed as homozygous *fie* plants (tables 4.6 to 4.8). Unfortunately, it is difficult from these results to compare proportions of morphologically abnormal seedlings from the three crosses as only the embryo specific family is homozygous for both transgenes.

The *fie*-2 allele contains a base change from G to A at the border between the 2nd intron and 3rd exon (Chaudhury A, personal communication). This mutation abolishes a native *Bsg*I recognition sequence, allowing for genotyping using CAPS markers (figure 4.16). The 'rapint1' and 'rapint5' introgenic primers pair were designed to amplify the genomic *FIE* sequence (and not the transgenic *FIE* sequence). The PCR fragments were then subjected to restriction analysis with *Bsg*I.

Figures 4.17 to 4.19 show two-week old homozygous *fie* seedlings. The seedlings were much smaller than wild type, leaves did not form, and there was no regular phyllotaxy.

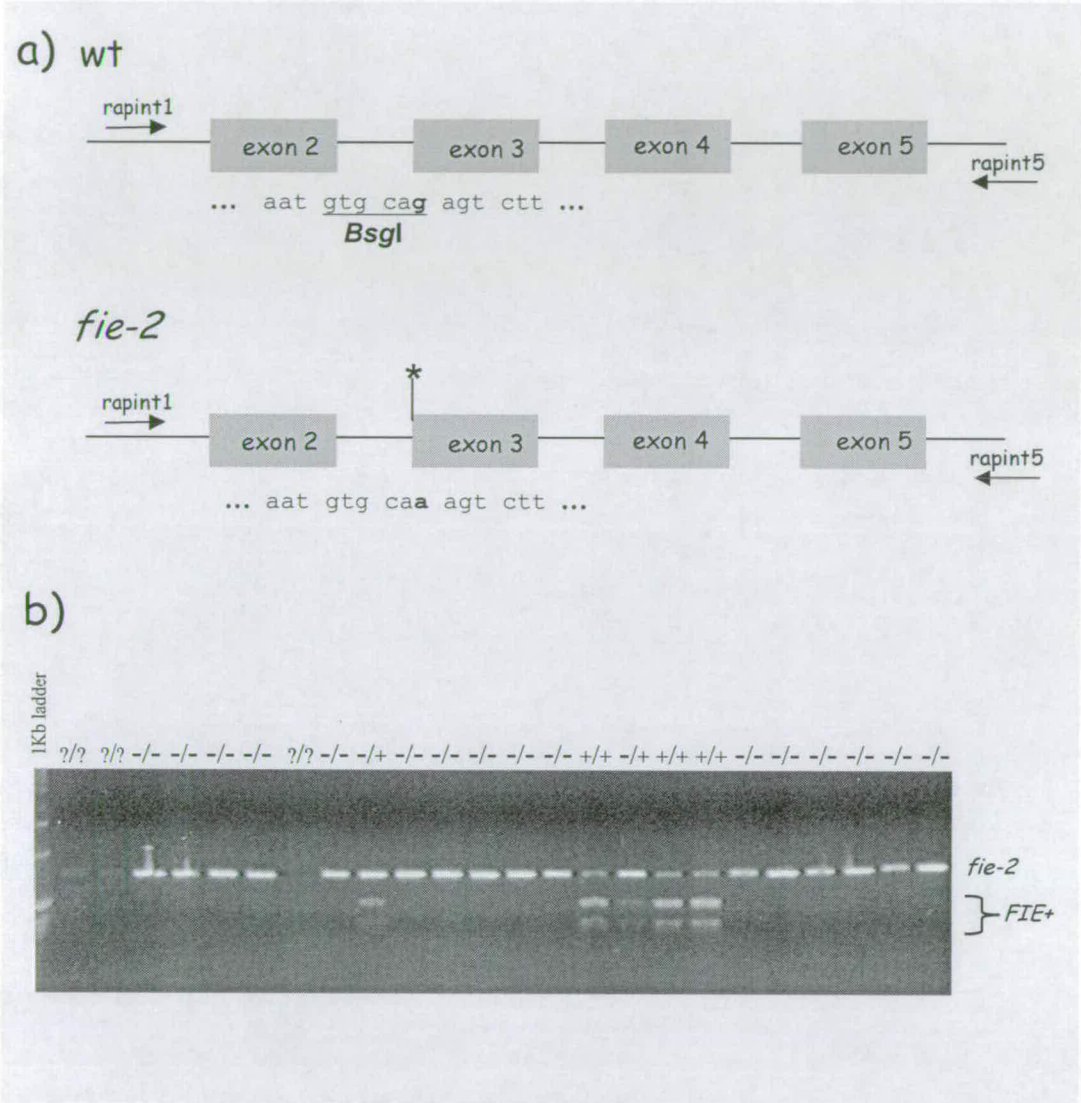


Figure 4.16 CAPS analysis to determine genotype at *FIE* locus. a) Sequence at *FIE* intron 2/exon 3 border showing position of *fie-2* lesion (red) and *Bsg*I restriction site. b) Gel showing PCR fragments after *Bsg* I digestion and deduced genotype. Note that enzyme does not cleave completely even for DNA from +/+ control plants. *fie-2* homozygotes can be distinguished as there is no cleavage whatsoever, and were confirmed by sequencing the PCR products.

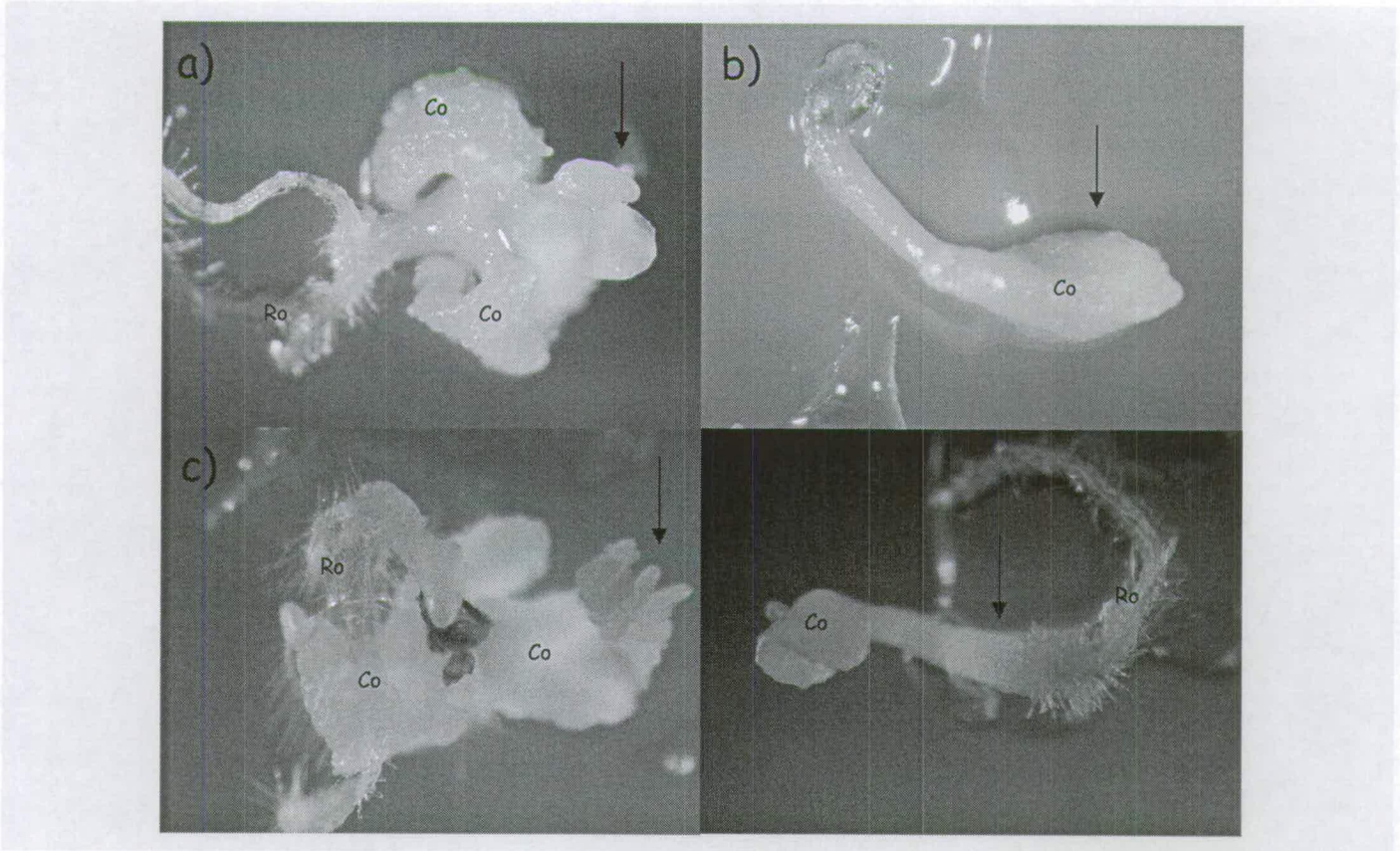


Figure 4.17 Phenotype of *fie* homozygotes from crosses involving embryo specific 9104 driver. a) Arrow indicates region of disorganized growth. b) Arrow indicates non-pigmented cotyledon. c) Arrow indicates repeatedly lobed region. d) Arrow indicates swollen hypocotyl. Co = Cotyledon, Ro = root.

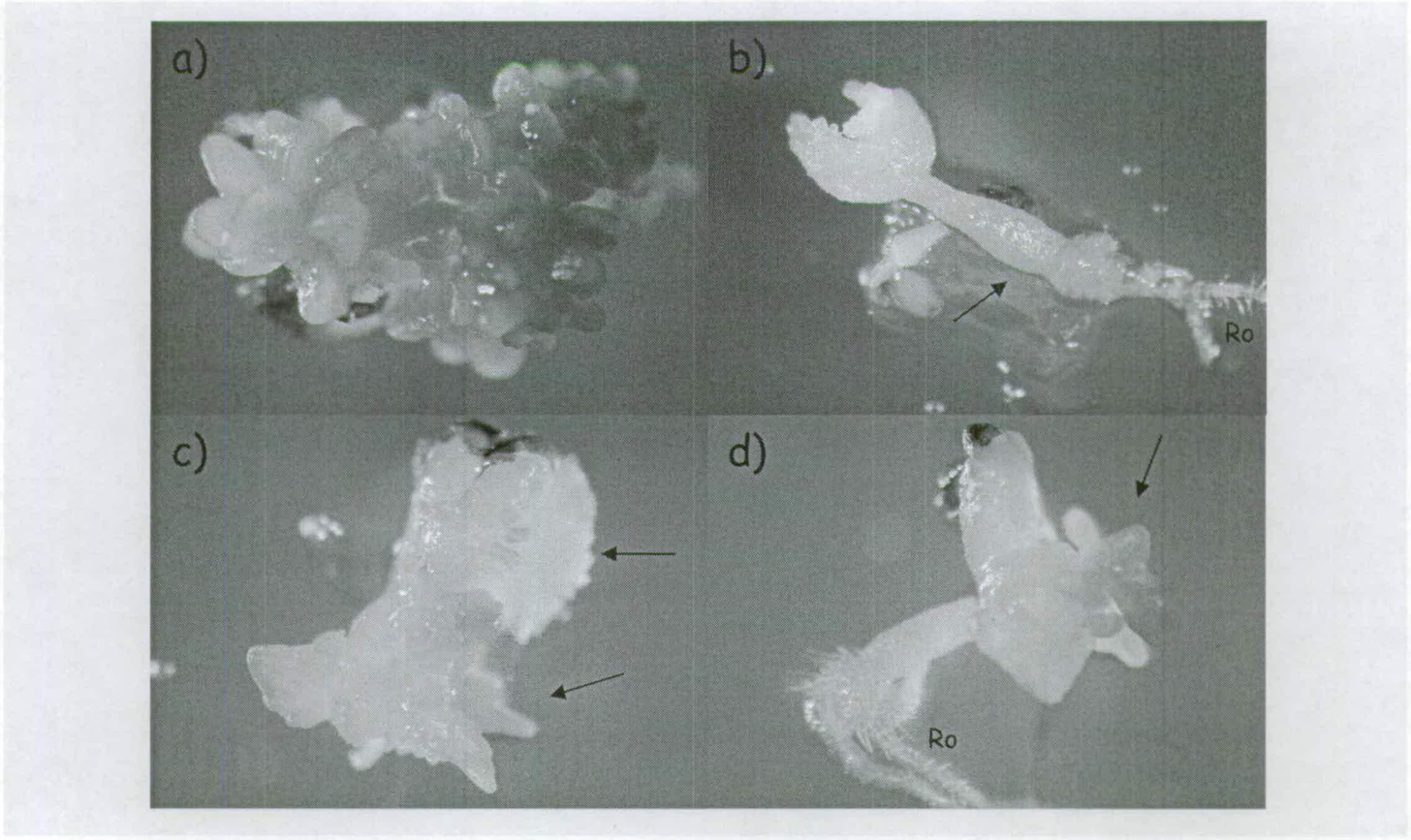


Figure 4.18 Phenotype of *fie* homozygotes from crosses involving early endosperm specific k22 driver. a) Repeatedly lobed. b) Arrow indicates swollen hypocotyl c) Arrows show region of disorganized growth. d) Arrow shows region of disorganized growth. Ro = roots.

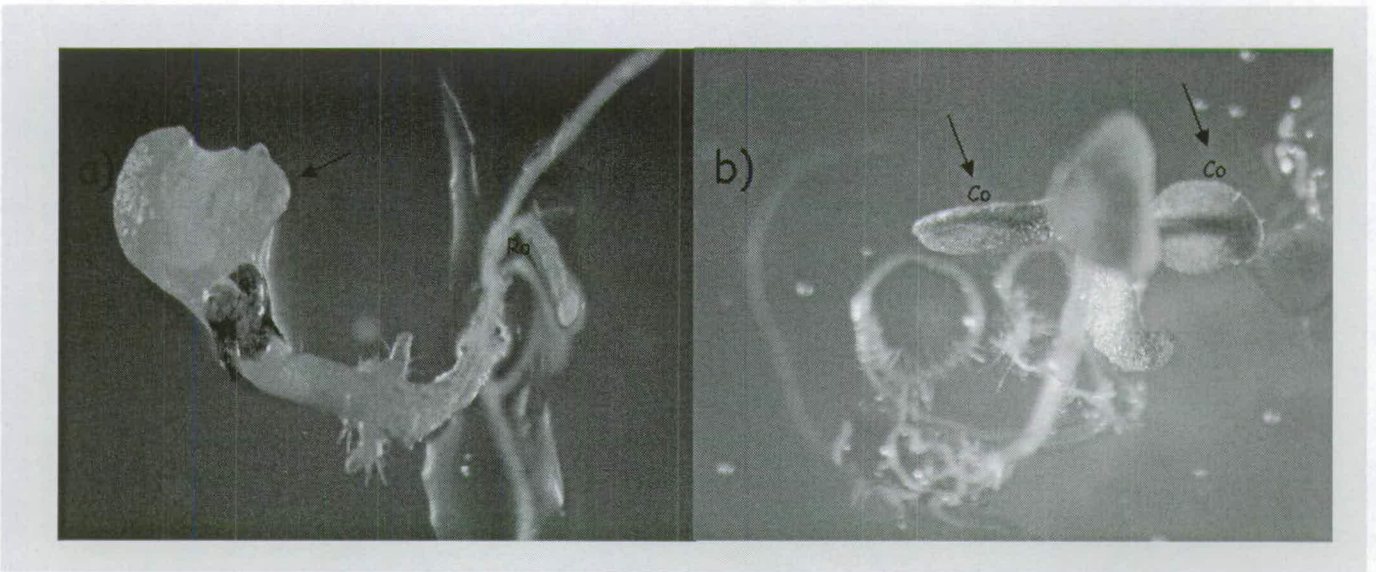


Figure 4.19 Phenotype of *fie* homozygote and *FIE+* seedling from crosses involving late endosperm specific 9185 driver. a) *fie-2* homozygous seedling. Arrow indicates abnormal cotyledon. b) *FIE+* seedling. Arrows indicate curled cotyledons. Co = Cotyledon, Ro = root.

Organs that did emerge were not well differentiated and often showed regions of disorganized growth or repeatedly lobed areas. Some seedlings possessed swollen hypocotyls or were etiolated. Occasionally, in some of the crosses, seedlings possessing curled cotyledons or leaves were observed (figure 4.19b) but CAPS and sequence analysis revealed that these seedlings were either heterozygous or wild type for *FIE*.

Family number	Proportion of seedlings appearing *fie-*
4/5	3.7% (7/188)
4/22	0.4% (1/262)
4/31	0.3% (1/354)
4/33	0.2% (1/416)
4/43	2.2% (11/496)
4/44	0.4% (2/456)
4/48	0.0% (0/176)

Table 4. 6 Proportion of morphologically abnormal seedlings of F4 seedlings homozygous for both embryo specific (9104) driver and slave transgenes. Actual numbers are bracketed.

Family Number	Proportion of seedlings appearing *fie-*
30/5	0.5% (1/200)
30/31	1.6% (4/240)
30/33	0.9% (4/440)
30/42	1.0% (4/420)

Table 4. 7 Proportion of morphologically abnormal seedlings of F3 seedlings involving the early endosperm specific (k22) driver. Actual numbers are bracketed. 30/5 and 30/33 are hemizygous for driver line insertion, 30/31 and 30/42 are homozygous. All plants were hemizygous for *UAS::FIE*.

Family number	Proportion of seedlings appearing *fie-*
230/1	0.0% (0/360)
230/4	0.0% (0/128)
230/5	0.0% (0/332)
230/30	0.4% (1/240)

Table 4. 8 Proportion of morphologically abnormal seedlings of F3 seedlings from crosses involving the late endosperm specific (9185) driver. Actual numbers are bracketed. 230/1 is homozygous for *UAS::FIE*. 230/4 and 230/5 contain no *UAS::FIE* insertion and 230/30 is hemizygous. All plants were hemizygous for driver trangene.

Kinoshita *et al.* (2001) was able to complement the *fie-1* embryo lethal phenotype by the introduction of a GFP fusion gene, *pFIE::FIE-GFP*, into *fie* mutants. However, the transgene expression did not recapitulate the wild type pattern of *FIE* RNA accumulation, perhaps because of a requirement for additional regulatory sequences. Transgenic

pFIE::FIE-GFP plants showed GFP fluorescence in the central cell nucleus before and after fertilization, which disappeared by the eight-nuclei endosperm, and could not be detected at any later stage of plant development. Rescued *fie* homozygotes developed into abnormal seedlings that were smaller than wild type and did not produce normal rosettes. In some cases highly disorganized structures emerged from the shoot apical meristem, and seedlings often produced floral-type organs. Kinoshita *et al.* conclude that the transition from vegetative to inflorescence development has prematurely occurred in the shoot apical meristem of seedlings homozygous for the *fie-1* allele and the *pFIE::FIE-GFP* transgene, and that this is reminiscent of *emf* mutants that skip the normal vegetative phase and flower early. Kinoshita *et al.* also reported flower-like structures generated from the hypocotyl and root, organs not usually associated with reproduction. However, no such flower-like structures and organs were observed in any of the *fie* homzygotes in this study. This difference could be due the different *fie* alleles used (although this is unlikely as both are thought to be null) or ecotypes used, or alternatively due the longer period of *FIE* expression in the developing seed from the *UAS::FIE* transgene used compared with the *pFIE::FIE-GFP* transgene. Furthermore, the *UAS::FIE* transgene may be active in the seedlings, allowing for some expression of FIE protein. The seedlings carrying each of the three driver line transgenes were therefore examined for GFP expression (figures 4.20 to 4.22) and compared with the autofluorescence of wild type Columbia seedlings (figure 4.23). Driver line 9104 showed expression throughout the vasculature of the seedling (figure 4.20). Driver line k22 showed expression in the epidermal cells of the leaves and cotyledons, the hypocotyl and root. Fluorescent ER bodies were also seen throughout the seedling (figure 4.21). Driver line 9185 showed no GFP expression in the leaves or cotyledons, but strong expression in the root, and weaker expression in the stem and petiole vasculature (figure 4.22). However, as none of the wild type-looking seedlings from the three crosses were identified as *fie* homozyotes it is likely that this limited expression pattern is not sufficient to recapitulate wild type *FIE* expression.

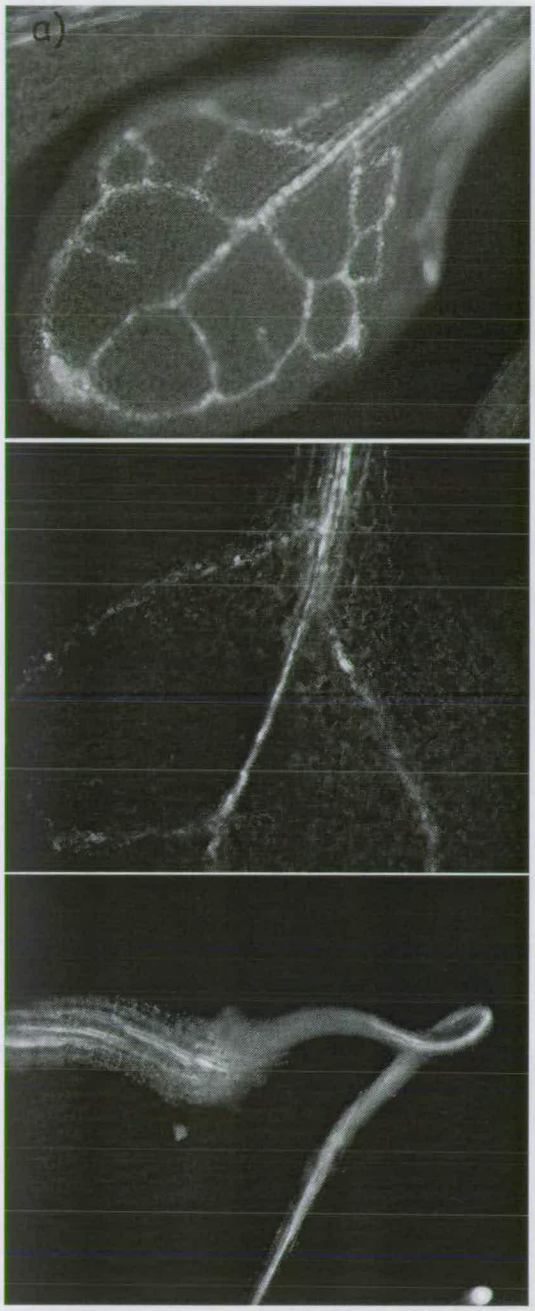


Figure 4.20 Seedling GFP expression in driver line 9104. Fluorescent microscopy. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. a) Leaf (red emission filter on). b) Cotyledon (red emission filter on). c) Root / hypocotyl junction.

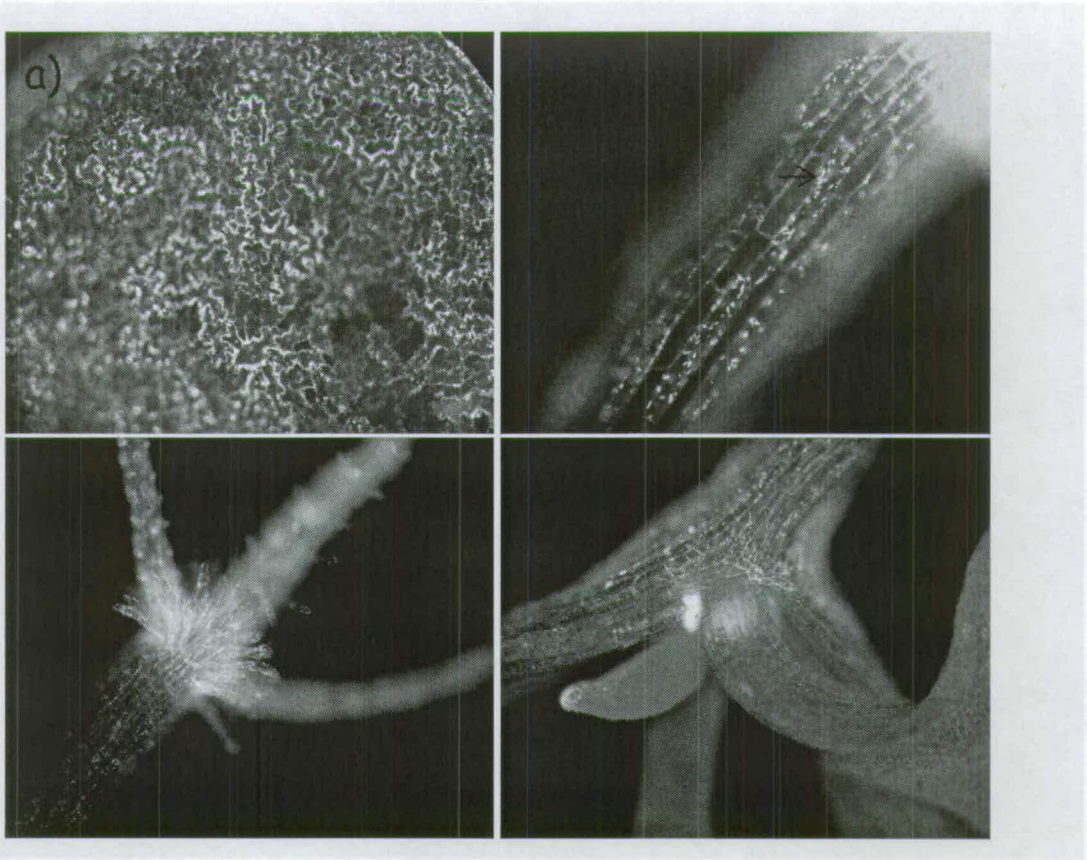


Figure 4.21 Seedling GFP expression in driver line k22. Fluorescent microscopy. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. a) Leaf (red emission filter on). b) Hypocotyl (red emission filter on). Arrow indicates an ER body. c) Hypocotyl/root junction. d) SAM.

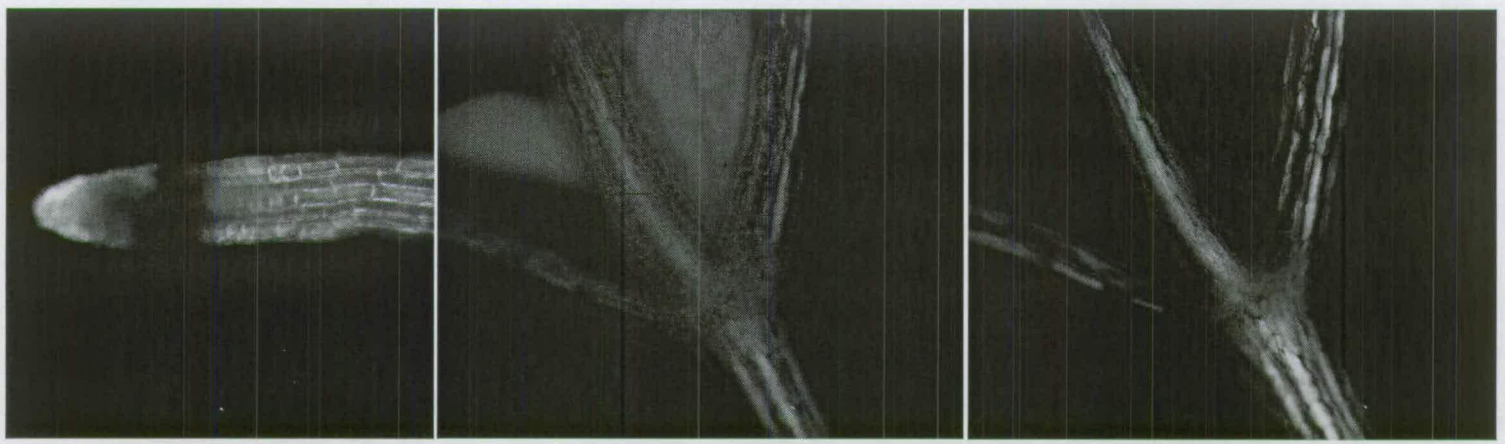


Figure 4.22 Seedling GFP expression in driver line 9185. Fluorescent microscopy. GFP fluorescence is seen as green, chlorophyll autofluorescence is seen as red. a) Root. b) and c) SAM. (c) red emission filter on).

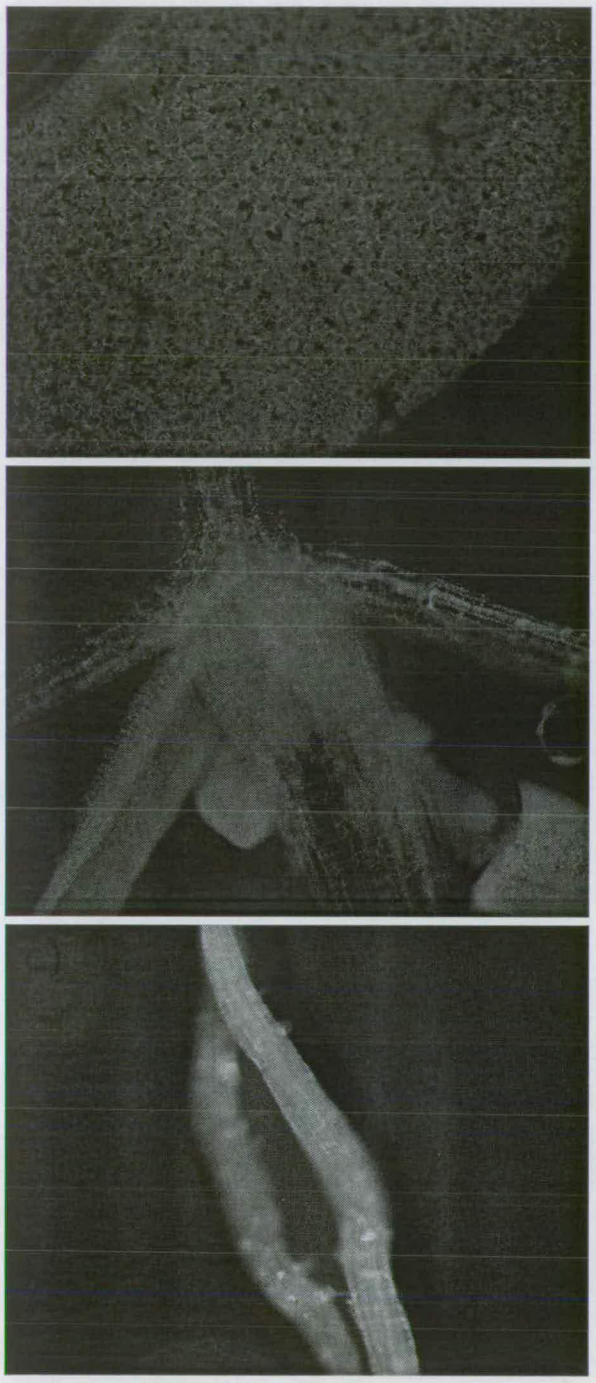


Figure 4.23 Wild type seedling (Columbia) autofluorescence. Chlorophyll autofluorescence is seen as red. a) Leaf. b) SAM. c) Root, showing a weak green autofluoresence.

Chapter 5. Discussion

5.1 A FIS Polycomb complex

5.2 A novel FIE interacting protein

5.3 The requirement for *FIE* activity in the embryo and endosperm

Chapter 5. Discussion

The results of this study show that MEA and FIE, and MEA and FIS2 interact in yeast two-hybrid, and that the interaction of MEA and FIE is mediated by the amino-terminal region of MEA, whilst interaction of MEA and FIS2 requires the FIS2 C–terminal VEFS box. Furthermore, as FIE, MEA, and FIS2 do not possess specific DNA binding activity, it remains unclear how the FIE complex is directed to its target genes. Yeast two-hybrid assays were used to identify a novel FIE interacting protein: RAP-1, a member of the basic helix-loop-helix (bHLH) family of transcription factors. However, the biological relevance of this interaction is unclear, as no genetic interaction of *FIE* and *RAP-1* is observed *in vivo*. Although, this could be due to redundancy as *RAP-1* is a member of a large gene family.

Although *FIE* is expressed in both embryo and endosperm, it is the endosperm that is predominantly affected in *fie* mutants. In order to establish if *FIE* is required in both tissues, I used a strategy for targeted expression of *FIE* to the embryo and endosperm regions in a *fie* background. I observed a rescue of *fie* embryo lethality so that viable *fie* homozygous seedlings were obtained. This confirms that the transactivation strategy used in this study is potentially valuable for dissecting *fie* function. However, it was not possible to resolve the site of *fie* function due to technical problems with one of the transgenic lines used.

5.1 A FIS Polycomb complex

Loss of FIE and MEA confers similar mutant phenotypes in both unfertilized embryo sacs and developing seeds; that is, fertilization-independent endosperm proliferation and the abortion of fertilized seeds that carry a maternally inherited mutant *mea* or *fie* allele (described in section 1.4). The shared mutant phenotype could be explained by the three *FIS* wild type gene products acting in a common pathway, whereby one gene activates the

expression of another to regulate seed development. Alternatively, the three FIS proteins may act as part of a multimeric complex similar to that observed with their animal homologues (section 1.5). This study used yeast two-hybrid to show that MEA and FIE interact, and that the *fie-6* mutation which drastically reduces *FIE* activity *in planta* also eliminates interaction of FIE and MEA, consistent with the FIE-MEA interaction being necessary for FIE activity *in planta*. The esc^{21} mimic mutation S255F (S301F in WD5 of the *Drosophila* protein ESC) introduced into the FIE bait (section 3.3) showed that it only partially disrupted its interaction with MEA in yeast (figure 3.5). This suggests that the S255F mutation might be less important in FIE compared with ESC for the interaction with its SET domain partner. It would be interesting to see if this was reflected in the trypsin sensitivity of the mutant FIE protein. Moreover, the S255F mutation could be engineered in a *FIE* transgene and introduced into a *fie* background. This might allow a weak mutant phenotype to be observed if the mutant FIE protein retained some ability to interact with MEA.

In situ data (Spillane *et al.* 2000 and Vielle-Calzada *et al.* 1999) show that *FIE* and *MEA* RNA expression overlaps *in vivo*, and cotransfection experiments (section 3.6) showed that the proteins co-localize within the nucleus *in planta* consistent with a physical interaction between these two proteins. In yeast, interaction of MEA and FIE was shown to be mediated by the amino-terminal region of MEA. Despite primary sequence divergence in this domain, MEA can interact with its corresponding animal partner ESC in the yeast two-hybrid system (figure 3.4). This is consistent with the secondary structure prediction of these regions. Both E(z) (Tie *et al.* 1998) and MEA are predicted to share some secondary structure and fold into a α-helix (figure 5.1).

The similarity of *fis2* mutants to *fie* and *mea* suggested that FIS2 might interact with either FIE, MEA, or both. However as discussed previously (section 3.4), no interaction with the

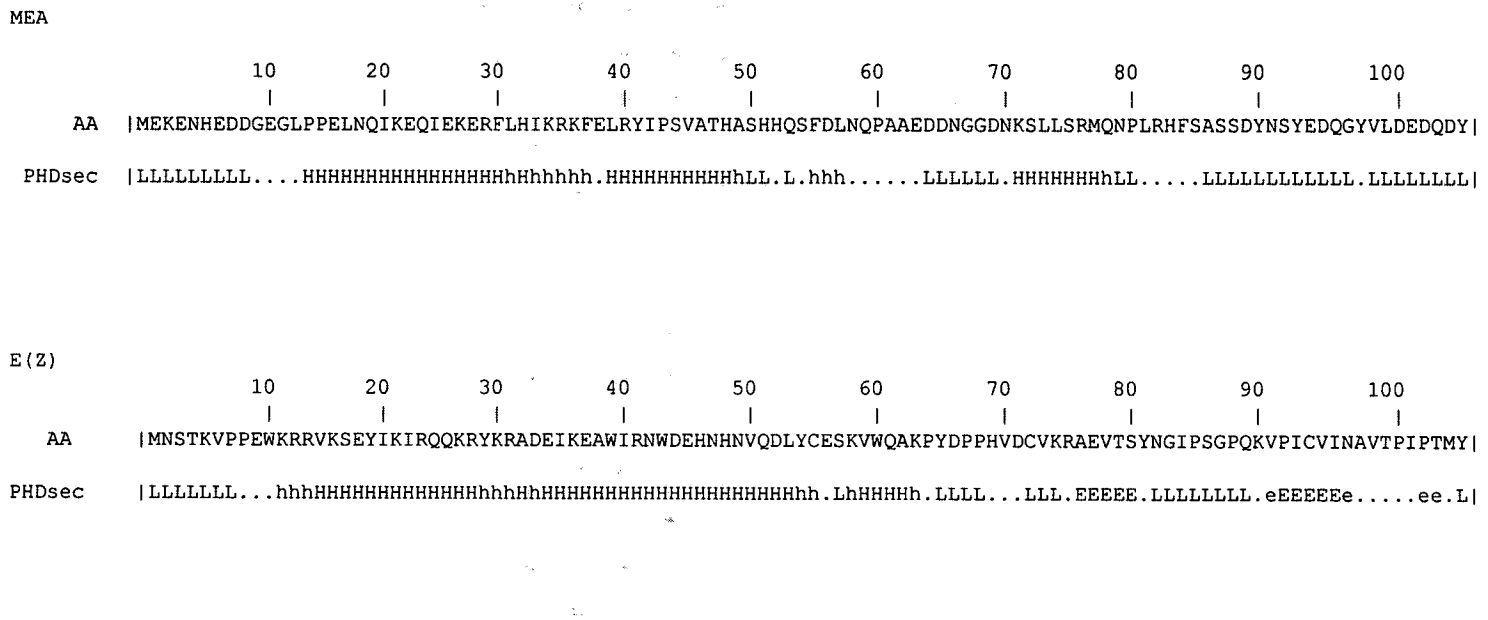


Figure 5.1 MEA and E(Z) N-terminal Protein Secondary Structure Prediction. Obtained using the NNPREDICT program. Both regions are predicted to include an α-helical structure. Prediction confidence: L >69% loop, H >82% helix, h <82% helix, E >82% extended (sheet), e <82% extended (sheet), - no prediction.

full-length FIS2 protein was observed in yeast two or three-hybrid, but that work by A. Bishopp (ICMB, Edinburgh) had suggested a possible reason for this. He found that the EMF2 protein, which is a homologue of FIS2, did not interact in yeast two-hybrid with the MEA homologue CLF when it was expressed as a full-length protein. However, a truncation containing only the C-terminus of EMF2, which contains the conserved VEFS box region did interact (A. Bishopp, unpublished). This might be attributed to steric hindrance of the full-length protein, whereby the interaction domain of the protein is restricted or obstructed in the chimeric fusion protein. Alternatively, the full-length FIS2 protein may be poorly expressed in yeast. The latter possibility is supported by the observation that both VRN2 and SuZ12, two homologues of FIS2, cannot be expressed as full-length proteins in *E.coli* (C. Dean, personal communication). The C-terminus of FIS2 was therefore expressed as a DB fusion, and tested for interaction with MEA. The results showed that MEA and the FIS2 VEFS box can interact in yeast, and that the FIS2 interaction domain is separate from the FIE interaction domain. No interaction with FIE was observed with either full-length FIS2 or the C-terminal VEFS box.

In vitro binding and co-immunoprecipitation experiments by Köhler *et al.* 2003 showed a direct physical interaction between FIE and MSI1, and that FIE, MEA and MSI1 are present together in a 600 kDa complex. Taken together these results suggest that the similarity of the *fie, mea, fis2,* and *msi1* phenotypes probably reflects the interaction of their gene products as part of a multiprotein complex (figure 5.2). By analogy with animal systems (van der Vlag J and Otte AP 1999), it is likely that this complex modifies histones to confer a transcriptionally inactive state.

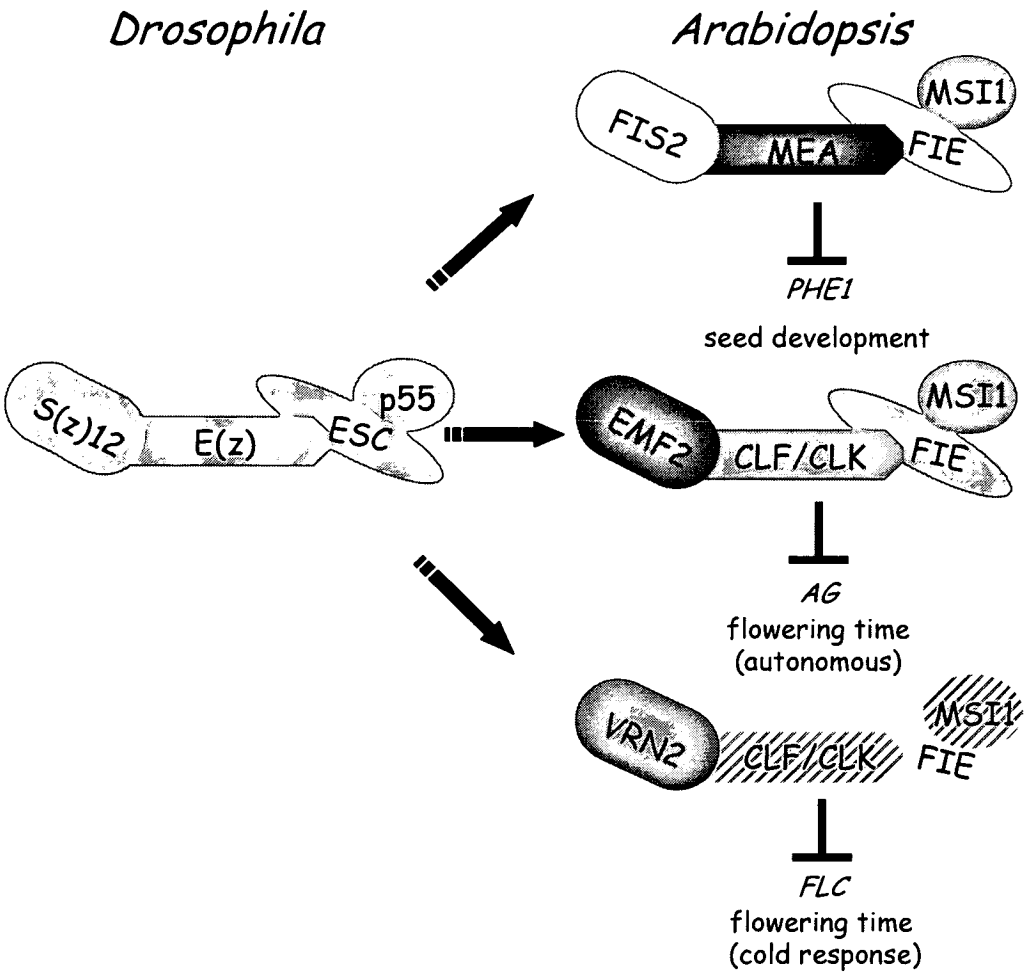


Figure 5.2 Diversification of PRC2 complexes in *Arabidopsis*. Three distinct complexes act throughout *Arabidopsis* development to repress target genes. Stripped proteins are not experimentally confirmed as complex members. See text for details.

5.1.1 Distinct Pc-G complexes in plants

In *Drosophila*, the PRC2 genes (including *ESC*, *E(Z)* and *Su(z)12*) are single copy genes. This is in contrast to *Arabidopsis*, where there is increasing support for the existence of at least three distinct plant PRC2 complexes. Yeast two-hybrid studies show that the Pc-G proteins CLF and CLK interact with FIE, and that EMF2 interacts with CLF (Y. Chanvivattana and A. Bishopp personal communication). The interaction of CLF with EMF2 in yeast was also confirmed by *in vitro* pull down experiments (D. Schubert personal communication). Consistent with a potential interaction *in planta*, CLF, CLK and EMF2 share overlapping expression domains by way of ubiquitous expression throughout the vegetative, inflorescence and floral meristems (Goodrich *et al.*, 1997 and Yoshida *et al.*, 2001). Further support for a functional physical interaction between these proteins is provided by the results of the phenotypic analysis of the various *Arabidopsis* Pc-G mutants. The *clf, emf2, fie* homozygotes, and *clf clk1-1* mutant plants exhibit overlapping phenotypes, and weak *clf* and *emf2* alleles show a strong genetic interaction (Yang *et al.*, 1995, Kinoshita *et al.*, 2001, and Y. Chanvivattana, personal communication). Furthermore, molecular analysis of partial loss-of-function *msi1* mutants revealed that MSI1 is required to maintain the correct temporal and organ-specific expression of the homeotic gene *AG,* a proposed target of CLF/CLK (Hennig *et al.* 2003). Taken together these results suggest that a distinct Pc-G complex containing CLF/CLK, FIE and EMF2 acts to repress the transition to flowering during vegetative development (figure 5.2).

A third plant Pc-G complex containing VRN2 is thought to epigenetically maintain the repression of *FLC* during vernalisation (Gendall *et al.*, 2001). It has recently been shown that VRN2 interacts with CLF and CLK in yeast (D. Schubert and A. Bishopp, personal communication). The complex is also likely to contain FIE, as other plant or animal E(z)-type proteins have always been found to act together with an ESC-type protein. Furthermore, vernalisation involves changes to the histone methylation status at the *FLC*

locus, increasing dimethylation of lysines 9 and 27 of histone H3 (Bastow *et al.* 2004). Such modifications are catalysed in animals by Pc-G complexes to maintain silenced chromatin states. Taken together these results suggest a functionally distinct Pc-G complex containing VRN2, FIE and CLF/CLK (figure 5.2).

Plants and animals evolved multicellularity independently of each other. The adoption of Pc-G-mediated repression to counteract the common problem of the maintenance of established developmental gene expression patterns through subsequent cell divisions during later development was, therefore, likely to have taken place independently between the two kingdoms. The ancestor sequence of the *ESC* and *E(z)* genes probably existed prior to the divergence of the two kingdoms. Later on, the polyploidisation events in plants resulted in most of the plant genes being found in multicopies, and due to selective pressure many of the duplicated genes acquired novel functions and expression patterns to act in diverse developmental pathways in the plant (figure 5.2). Nevertheless, some redundancy is still thought to remain, for example, between *CLF* and *CLK* (Y. Chanvivattana, personal communication). Also, as the *fie* seedling phenotype and *clf clk* phenotype is stronger than the *emf2* phenotype it is likely that *EMF2* might prove to be partially redundant with the ubiquitously expressed *VRN2*. Accordingly, it would be interesting to see if the *emf2 vrn2* double mutant shows an enhanced phenotype.

As previously mentioned (in section 1.5.1), at least two biochemically distinct *Drosophila* Pc-G complexes have been identified. The ESC-E(z) complex (PRC2) acts early to establish silencing, while the PRC1 complex is recruited later and functions to maintain silencing throughout development (Shao *et al.*, 1999, Ng *et al.*, and Tie *et al.*, 2001). Interestingly, no *Arabidopsis* homologues to the PRC1 Pc-G proteins have yet been identified. This could be considered consistent with the idea that plant development is much more flexible than animal development, and *Arabidopsis* does not possess an equivalent long-term memory system.

Alternatively, *Arabidopsis* could have as yet unidentified proteins equivalent in function to the animal PRC1 that interpret and maintain epigenetic marks, such as methylated K27 on histone H3, but which can be easily erased by other enzymic systems that are more widely expressed in plants.

5.2 A novel FIE interacting protein

In *Drosophila* it remains unknown how the ESC-E(z) complex is recruited to its target genes. The transcription factor Hunchback (Hb) has been implicated, but a physical interaction with the complex has not yet been demonstrated (Kehle *et al.*, 1998). However, it has recently been confirmed that several short motifs are required for PRE/TRE function, these include binding sites for three sequence-specific DNA binding proteins: the Pc-G protein Pleiohomeotic (PHO), and the GAGA factor and zeste protein, both of which are trxG members, suggesting that the Pc-G and trxG complexes are recruited to their target genes by these sequence specific factors (Ringrose *et al.* 2003).

As discussed previously, the *Arabidopsis* FIS proteins (FIE, MEA and FIS2) and MSII act together in a complex to repress its target gene(s) (*e.g. PHE1*). However, as none of these proteins have been found to possess specific DNA-binding activity it remains unclear how this complex is targeted. It is therefore thought that the FIS complex could contain proteins that possess specific DNA-binding ability to target the complex to its DNA target gene(s). The yeast two-hybrid screen carried out using FIE-DB as 'bait' identified one FIE interacting protein: R-homologous *Arabidopsis* Protein-1 (RAP-1), a member of the basic helix-loop-helix (bHLH) superfamily of myc-like transcription factors (figure 3.11a). A disadvantage of the yeast two-hybrid system is the occurrence of false positive interactions, so to help eliminate the possibly of this being the case for FIE and RAP-1, *in vitro* pull-down

experiments could be performed. Co-immunoprecipitation experiments could be performed to confirm the interaction *in planta*.

5.2.1 The basic helix-loop-helix family

RAP-1 is a member of the basic helix-loop-helix (bHLH) superfamily of myc-like transcription factors. It is thought that the bHLH domain is an ancient component of transcriptional regulation as it has been found not only in the plant kingdom, but also in animals and in yeast. bHLH proteins have been most widely studied in mammalian and other non-plant eukaryotic systems where they are found to control a diversity of processes such as cell proliferation, neurogenesis, sex determination, and cell fate specification (Grandori *et al.* 2000, Massari and Murre 2000). Members of this superfamily are characterised by a bipartite bHLH domain, consisting of ~60 amino acids. At the N terminal end of the domain lies a stretch of ~15 amino acids with a high number of basic residues. The rest of the domain consists of the HLH made up of mainly hydrophobic residues that form two amphipathic α-helices separated by a loop region (figure 3.11b). The HLH region allows dimerisation to occur allowing the formation of homo and hetrodimers. This association brings together the two basic regions that then bind to the DNA recognition sequence. The core of the DNA sequence motif recognised by the bHLH proteins is a six-nucleotide stretch of DNA known as the E-box (5'-CANNTG-3'). The ability to heterodimerise with other bHLH partners allows for a large number of potential DNA-binding complexes, each with variable DNA binding properties and activation or repression abilities. The potential is further increased by the spatial expression patterns of the heterodimer components or MYB partners.

The *Zea mays* R gene product Lc (involved in anthocyanin synthesis) was the first plant bHLH protein to be discovered (Ludwig *et al.* 1989). Since then, between 118 and 147 bHLH genes have been identified in *Arabidopsis* (Buck and Atchley 2003, Heim *et al.* 2003,

Toledo-Ortiz *et al.* 2003), but only a few of these have been characterised in detail. Those that have been characterised have been found to function in a wide range of processes including anthocyanin biosynthesis, phytochrome signalling, globulin expression, fruit dehiscence, carpel, and epidermal development. Although the bHLH region of these proteins is well conserved in the plant superfamily the remainder of the protein is generally much less conserved. So far, no close animal homologues of the plant bHLH proteins have been found. Using just the bHLH domains for alignment Toledo-Ortiz *et al.* (2003) created a phylogenetic tree showing 21 subfamilies, whereas Heim *et al.* (2003) describe 25 subfamilies, and Buck and Atchley (2003), 15. 10% of *Arabidopsis* bHLH genes do not possess introns (Heim *et al.* 2003), including *RAP-1* and its close homologues *bHLH28* and *bHLH5/ATR2* (Smolen *et al.* 2002). One reason for this could be retrotransposition: a process whereby genes are duplicated by using mature mRNA as the template. Intronic and gene distributions throughout the *Arabidopsis* genome suggests that overall, 38% of bHLH genes could have evolved by some type of duplication event (Heim *et al.* 2003, Toledo-Ortiz *et al.* 2003), and unlike in animals, some plant bHLH genes are partially redundant (Goff *et al.* 1992, Riechmann and Ratcliffe 2000).

5.2.2 R-homologous *Arabidopsis* Protein-1 (RAP-1)

RAP-1 was first identified in a southwestern screen by de Pater *et al.* (1997). The screen was designed to identify *Arabidopsis* proteins that could bind to elements derived from the pea lectin promoter. Further investigation showed that RAP-1 protein binds to the G –box sequence 5'-CACNTG-3'. The expression pattern of *rap-1* analyzed by northern blot hybridization showed that *rap-1* is expressed mainly in vegetative tissues. The highest expression was observed in the root and stem, and no expression was found in siliques. *rap-1* was shown to be functionally different from R-proteins such as Lc.

Abe *et al.* (1997) identified *RAP-1* independently in a separate binding screen, using a fragment of the *RD22* promoter containing two G-boxes (5'-CACATG-3'). *RD22* is an *Arabidopsis* dehydration–responsive gene (Yamaguchi-Shinozaki and Shinozaki 1993). Abe *et al.* (1997) refer to *RAP-1* as *RD22BP1* (*RD22* binding protein 1), and later as AtMYC2 (Abe *et al.* 2003), however for the sake of precedence the name *RAP-1* will be retained here. RNA gel blot analysis revealed that transcription of *RAP-1* is induced by dehydration stress and abscisic acid (ABA) treatment, and its induction precedes that of *RD22*. During drought conditions plant cells sense the loss of water, transducing a stress signal to the nuclei. This leads to the expression of specific drought response genes. Most of these genes are also induced by ABA, although there are also ABA-independent drought responsive genes. *RAP-1* itself is rapidly induced by ABA treatment, dehydration, and high salt treatment. No increase in *RAP-1* mRNA was observed in cold stressed plants (Abe *et al.* 1997).

Northern-blot analysis carried out by Abe *et al.* (1997) showed a different pattern to that previously described by Pater *et al.* (1997), with strong expression in siliques and stems, but weak in leaves and roots. Transgenic plants over-expressing *RAP-1* show hypersensitivity to ABA and increased levels of *RD22* and *AtADH1* (*Alcohol Dehydrogenase 1*) expression (Abe *et al.* 2003). *Rap-1* loss-of-function lines were generated using an Ac/Ds tagging system and were found to exhibit reduced ABA sensitivity and showed decreased expression levels of *RD22* and *AtADH1* upon ABA treatment compared to wild type, and were significantly retarded in growth (Abe *et al.* 2003). The *RD22* promoter region also contains a MYB2 binding site, and it is thought that together with MYB2, RAP-1 activates *RD22* expression in drought conditions (Abe *et al.* 2003).

5.2.3 The *rap-1* mutant and genetic interactions with Pc-G members

To determine whether the interaction of RAP-1 and FIE in yeast might be biologically relevant *in planta*, plants carrying two separate *rap-1* insertion alleles were first identified

using the Salk T-DNA insertion collection. The T-DNA insertions were confirmed and predicted to cause severe loss-of-function or null mutations. Subsequently, *rap-1* plants were examined for phenotype such as those observed with *fie* mutants and crossed to *fie* and *clf* to look for genetic interactions in double mutant combinations. However, no apparent *rap-1* phenotype was observed (figure 3.16). Because FIE is likely to act with other Pc-G members after germination as discussed above, whole plant phenotypes (for example, similarities to *clf* mutants) were also characterised. However, no apparent vegetative or floral mutant phenotype was observed (data not shown). Abe *et al.* 2003 also reported no observable phenotype unless the plants were first treated with ABA. In this case growth retardation was reduced significantly in ABA treated plants compared with wild type plants consistent with WT *RAP-1* inducing the stress response gene *RD22* (Abe *et al.* 2003). It would be interesting to look for any similar ABA sensitivity in *fie*, *rap-1 fie*, *clf/clk*, *vrn2* or *emf2* plants, as this would support the biological relevance of the FIE-RAP-1 physical interaction.

To test for a genetic interaction of *RAP-1* with *FIE in planta*, *rap-1* mutants were introduced into a *fie* background to assess whether the *fie* phenotype was modified. If RAP-1 acts with FIE in a complex one might expect to see a genetic interaction, for example enhancement of seed or vegetative phenotype. However, double mutant plants could not be generated as homozygous *fie-2* plants abort as embryos. Nonetheless, a reduced dosage of *FIE* might have an effect in a genetic background that is already compromised for a different member of the complex, *e.g. RAP-1*. However, no differences from the characteristic *fie* phenotype were observed. The vegetative phenotype of these plants was also examined but no mutant phenotype was seen. Nevertheless, these results do not exclude a genetic interaction between *FIE* and *RAP-1* for several reasons: members of a common complex do not always show dosage effects, *RAP-1* may be redundant, and the *rap-1* phenotype may be subtle or not apparent with the growth conditions used.

As RAP-1 interacts with FIE in yeast, and FIE is thought to be a component of several distinct Pc-G complexes, there is a potential for genetic interaction between *RAP-1* and *CLF* or *VRN2 in planta*. *rap-1* mutants were therefore crossed into the *clf* background. However, no enhancement of the *clf* phenotype was observed. Although, again this lack of genetic interaction could be attributed to a functional redundancy of the *RAP-1* gene. As FIE is also thought to be a component of a complex containing VRN2, I crossed *rap-1* into two late flowering backgrounds in which the response to vernalisation treatment is most pronounced, unfortunately due to time constraints the F2 seed were not vernalised for a sufficient duration and a clear response to the cold treatment was not seen, and the results were therefore inconclusive (data not shown). Together these results show that FIE and RAP-1 proteins have the potential to interact in yeast two-hybrid, but it remains unclear if this interaction is biologically relevant *in planta*, and whether RAP-1 is involved in recruiting any of the FIE containing complexes to their target genes.

5.3 The requirement for *FIE* activity in the embryo and endosperm

As previously discussed, *FIE* expression analysis (demonstrated by *in situ* hybridisation) shows that it is expressed in both embryo and endosperm during early seed development (Spillane *et al.* 2000). The distribution of *FIE* mRNA in developing seed is consistent with FIE functioning in both early embryo and endosperm development after fertilization. However, although aborting *fie* seed show defects in both endosperm and embryo development, the embryo relies on the endosperm for nourishment and signals, raising the possibility that the embryo arrest could be an indirect consequence of defects in endosperm development. As previously discussed (in section 4.10), Kinoshita *et al.* (2001) reported the rescue of *fie* lethality using a *pFIE::FIE-GFP* transgene that showed GFP expression in the endosperm up until the 8-cell stage, and no expression in the embryo at any stage, consistent with the requirement of *FIE* in the endosperm alone. However, there is some evidence to

suggest that *FIE* is also required in the embryo. Occasionally, *mea/mea* or *fis2/fis2* plants are obtained, but never *fie* homozygotes, suggesting *FIE* may have additional functions in the embryo, consistent with its expression there. Alternatively, as unlike *MEA* and *FIS2*, *FIE* is a single copy gene, so *mea* and *fis2* may just have a weaker phenotype due to redundancy, for example with *CLK* or *EMF2* respectively. In order to establish the requirements for *FIE* in the embryo and endosperm I targeted *FIE* expression to specific domains of the seed, using a driver-slave strategy (summarised in figure 4.1). Three driver lines were used expressing the activator GAL4-VP16 in the embryo (9184), the early endosperm (k22), or in the late endosperm (9185). Unfortunately, due to technical difficulties discussed previously (in section 4.6), only one *UAS::FIE* slave line was used (14.2).

Targeted expression of the *FIE* transgene specifically in the embryo or endosperm in a *fie* background could give rescue of the seed abortion phenotype. In this case *fie/+* plants homozygous for both driver and slave transgenes will show 100% seed viability. If heterozygous for the driver transgene, then it is likely that only seed inheriting the transgene maternally will be rescued, giving 25% seed abortion, whilst plants hemizygous for both transgenes would display around 37.5% seed abortion. Significantly lowered abortion ratios were obtained in all of the three crosses analysed. The F3 plants homozygous for both slave and embryo-specific driver transgenes showed some plants with abortion ratios as low as 40%. This suggests that some rescue of the *fie* embryo lethality was occurring. The fact that there was 40% seed aborting and not 0% might be because the *UAS::FIE* was not expressed in sufficient levels to give full complementation. The F2 plants segregating for both slave and early or late endosperm-specific driver transgenes showed some plants with abortion ratios as low as 29% and 38% respectively. Although it was not determined whether these plants are homozygous or hemizygous for the slave transgene, it is interesting to note that the plants with significantly lower abortion ratios were not always homozygous for the driver

transgene, as might be expected. Also, three plants showing no GFP expression (30/20, 30/46, and 230/47), although not exhibiting significantly reduced abortion ratios, displayed some abnormally large seed, thought to be associated with a *fie* rescue phenotype. This suggests that the 14.2 slave line transgene could be inserted in a context that allows for some basal expression of *UAS::FIE* regardless of GAL4-VP16 activation.

If the *fie*- phenotype of embryo abortion was due to an absolute requirement for *FIE* activity in either the embryo or endosperm, one might expect to see a difference in the proportion of plants displaying lower abortion ratios depending on the driver line used. For example, if *FIE* is not required in the endosperm, but only in the embryo, one would expect no rescue of embryo abortion in the crosses involving the endosperm-specific drivers, k22 and 9185. However, as some seed abortion percentages are lower (29%) in the early-endosperm specific cross than the other crosses (~39%), it suggests that *FIE* might be principally required in the early endosperm in wild type plants. Alternatively, as there was originally thought to be two unlinked insertions in slave line 14.2, it is possible that the early endosperm specific cross contains a *UAS::FIE* insertion at a separate locus from those in the other two crosses, resulting in an increased GAL4-independent expression. However, further crosses with independent *UAS::FIE* slave lines would need to be carried out to support this.

As discussed previously (in section 4.9), the novel phenotype of abnormally large, late desiccating seed is likely due to *fie* rescue, but it would be interesting to confirm this by sequence or CAPS analysis of the *FIE* locus in these seed. A similar strategy was used to analyse the *FIE* locus in the morphologically abnormal seedlings, which demonstrated that it was possible to obtain viable homozygous *fie* seedlings from all three crosses. However, it would be necessary to obtain lines homozygous for both transgenes to be able to compare the

frequency of *fie* homozygotes between the crosses, to establish if the early-endosperm specific lines are showing a higher frequency of rescue.

The delayed activity of a paternal copy of a *FIE+* allele is not sufficient to rescue a *fie* ovule. It is interesting therefore, that at least one *fie* homozygote seedling was identified from the cross involving the late endosperm-specific driver transgene. In this driver line (9185), transgene expression is not detected in the endosperm until the early torpedo stage of embryo development, a developmental stage later than at which paternal *FIE* expression begins. Taken together, it appears the results obtained are not very informative as to the required site of *FIE* activity in the developing seed. However, it is possible that *FIE* could be required in both the endosperm and embryo, but expression in any one of these tissues can rescue the lethality. Alternatively, it is possible that there is low level expression of the driver transgene in the endosperm/embryo that is too weak to be detected by confocal microscopy analysis of GFP fluorescence, but sufficient for some GAL4-VP16 activation of *UAS::FIE*.

The *fie* homozyotes rescued as seedlings were morphologically abnormal. The seedlings were much smaller than wild type, leaves did not form, and there was no regular phyllotaxy. Organs that did emerge were not well differentiated and often showed regions of disorganized growth or repeatedly lobed areas. Some seedlings possessed swollen hypocotyls or were etiolated. No floral organs or tissue were observed as reported by Kinoshita *et al.* (2001). It would therefore be interesting to test these *fie* plants for misexpression of floral homeotic genes also reported by Kinoshita *et al.* (2001). As it is thought that a FIE Pc-G complex is required to prevent early flowering during the vegetative phase, the absence of floral tissue in these seedlings could be due to the longer period of *FIE* expression in the developing seed from the *UAS::FIE* transgene used in this study compared with the *pFIE::FIE-GFP* transgene Kinosita *et al.* used. Or alternatively, that the seedling expression of the driver transgenes was sufficient to give some limited FIE activity.

Taken together these results show that the driver-slave strategy for *FIE* transactivation in the developing seed is capable of rescuing the *fie* embryo lethality, however, it remains unclear as to the requirement for *FIE* in the endosperm or embryo.

References

Abe, H *et al.* **Role of Arabidopsis MYC and MYB Homologs in Drought and Abscisic Acid-Regulated Gene Expression.** *Plant Cell.* **9:** 1859-1868 (1997).

Abe, H *et al.* **Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling [sic].** *Plant Cell.* **15:** 63-78 (2003).

Ach, RA *et al.* **A conserved family of WD-40 proteins binds to the retinoblastoma protein in both plants and animals.** *Plant Cell.* **9:** 1595-1606 (1997).

Adams, S *et al.* **Parent-of origin effects on seed development in Arabidopsis thaliana require DNA methylation.** *Development.* **127**:2493-2502 (2000).

Agrawal, AF and Chasnov, JR. **Recessive mutations and the maintenance of sex in structured populations.** *Genetics.* **158:**913-917 (2003).

Alonso, JM *et al.* **Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana*.** *Science* **301**: 653-657. (2003).

Altschul, SF *et al.* **Basic local alignment search tool.** *J. Mol. Biol.* **215**:403-410 (1990).

Aubert, D *et al.* **EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis.** *Plant Cell.* **13:**1865-1875 (2001).

Baroux, C *et al.* **Paternally inherited transgenes are down-regulated but retain low activity during early embryogenesis in *Arabidopsis*.** *FEBS Letters.* **509:11-16 (2001).**

Bartel, P. L. **Using the two hybrid system to detect protein-protein interactions**, p. 153–179. *In* P. L. Bartel, C.-T. Chien, R. Sternglanz, and S. Fields (ed.), Cellular interactions in development: a practical approach. Oxford University Press, Oxford (1993).

Baulcombe, DC *et al.* **Biologically-active viral satellite RNA from the nuclear genome of transformed plants.** *Nature*. **321(6068):** 446-449 (1986).

Becker, D. **Binary vectors which allow the exchange of plant selectable markers and reporter genes.** *Nucleic Acids Res.* **11;18(1)**:203 (1990).

Berger, F. **Endosperm Development.** *Curr. Opin. Plant Biol.* **2:**28-32 (1999).

Berger, F. **Endosperm: the crossroad of seed development**. *Curr. Opin. in Plant Biol.* **6:**42-50 (2003).

Birve, A *et al.* ***Su(z)12,* a novel *Drosophila* polycomb group gene that is conserved in vertebrates and plants.** *Development*. **128**:3371-3379 (2001).

Blanc, G *et al.* **Extensive gene duplication and reshuffling in the Arabidopsis genome**. *The Plant Cell.* **12:**1093-1101 (2000).

Boisnard-Lorig, C *et al.* **Dynamic analyses of the expression of the HISTONE::YFP fusion protein in *Arabidopsis* show that syncytial endosperm is divided in mitotic domains.** *Plant Cell.* **13:**495-509 (2001).

van Bokhoven, H *et al.* **Protoplasts transiently expressing the 200K coding sequence of cowpea mosaic virus B-RNA support replication of M-RNA.** *J. Gen. Virol.* **74:**2233-2241 (1993).

Bowman, JL *et al.* **Genes Directing Flower Development in *Arabidopsis.*** *The Plant Cell.* **1**: 37-52 (1989).

Brand, AH and Perrimon, N. **Targeted gene expression as a means of altering cell fates and generating dominant phenotypes.** *Development.* **118:** 401-415 (1993).

Buck, MJ and Atchely, WR. **Phylogenetic Analysis of Plant Basic Helix-Loop-Helix Proteins.** *J Mol Evol.* **56:**742-750 (2003).

Cao, R *et al.* **Role of Histone H3 lysine 27 methylation in Polycomb-group silencing.** *Science.* **298**:1039-1043 (2002).

Carpenter, R and Coen ES. **Floral homeotic mutations produced by transposon-mutagenesis in *Antirrhinum majus.*** *Genes Dev.* **4**:1483-93 (1990).

Chaudhury, AM *et al.* **Fertilization-independent seed development in *Arabidopsis thaliana.*** *Proc. Natl. Acad. Sci. USA.* **94:** 4223-4228. (1997).

Chaudhury, AM *et al.* **Ovule and Embryo development, apomixis and fertilization.** *Curr. Opin. in Plant Biol.* **1**:26-31. (1998).

Choi, Y *et al.* **DME, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis.*** *Cell.* **110:**33-42 (2002).

Chuang, CF and Meyerowitz, EM. **Specific and heritable genetic interference by double-stranded RNA in Arabidopsis thaliana.** *Proc. Natl. Acad. Sci. USA.* **97**:4985–4990. (2000).

Coen, ES and Meyerowitz, EM. **The war of the whorls: genetic interactions controlling flower development.** *Nature.* **353**:31-37 (1991).

Crews, D and Fitzgerald, KT. **Sexual behaviour in parthenogenetic lizards (Cnemidophorus).** *Proc. Natl. Acad. Sci. USA* **77**:499-502 (1980).

Czermin, B *et al.* ***Drosophila*** **Enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites.** *Cell.* **111**:185-196 (2002).

Dejong, AJ *et al.* **Early Events In Higher-Plant Embryogenesis.** *Plant Mol. Biol.* **22 (2):** 367-377 (1993).

Denisenko, O *et al.* **The Product of the Murine Homolog of the *Drosophila extra sex combs* Gene Displays Transcriptional Repressor Activity.** *Mol. Cell. Biol.* **17**: 4707-4717 (1997).

Denisenko, O *et al.* **Point mutations in the WD40 Domain of Eed Block Its Interaction with Ezh2.** *Mol. Cell. Biol.* **18**: 5634-5642 (1998).

Dickinson, H and Scott, R. **DEMETER, goddess of the harvest, activates maternal MEDEA to produce the perfect seed.** *Mol. Cell.* **10:** 5-7 (2002).

Drews, GN *et al.* **Genetic Analysis of Female Gametophyte Development and Function.** *The Plant Cell.* **10**: 5-17 (1998).

Dubreucq B, *et al.* **The Arabidopsis AtEPR1 extensin-like gene is specifically expressed in endosperm during seed germination.** *Plant J.* **23 (5):** 643-652 (2000).

Fields, S and Sternglanz, R. **The two-hybrid system: an assay for protein-protein interactions.** *Trends Genet.* **10(8):**286-92 (1994).

Fields, S and Song, OK. **A NOVEL GENETIC SYSTEM TO DETECT PROTEIN PROTEIN INTERACTIONS.** *Nature.* **340 (6230)**: 245-246 (1989).

Faure, JE *et al.* **Fertilization in *Arabidopsis thaliana* wild type: developmental stages and time course.** *Plant J.* **30(4):** 481-488 (2002).

Garcia, D *et al.* **Arabidopsis *haiku* Mutants Reveal New Controls of Seed Size by Endosperm.** *Plant Physiology.* **131**: 1661-1670 (2003).

Gendall, AR *et al.* **The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis.** *Cell.* **107 (4):** 525-535 (2001).

Goff, SA *et al.* **Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins.** *Genes & Devt.* **6:**864-875 (1992).

Golemis EA, *et al.* **Interaction trap/two-hybrid system to identify interacting proteins.** In Current Protocols in Molecular Biology. Edited by Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. John Wiley: New York; 1996:20.1.1–20-1-23.

Golemis, EA and Khazak, V. **Alternative yeast two-hybrid systems. The interaction trap and interaction mating.** *Methods Mol. Biol.* **63**, 197-218 (1997).

Goodrich, WJ *et al.* **A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*.** *Nature*. **386**: 44-51 (1997).

Goodrich, J. **Plant development: Medea's maternal instinct.** *Curr. Biol.* **8:** R480-R484. (1998).

Goodrich, J and Tweedie, S. **REMEMBRANCE OF THINGS PAST: Chromatin Remodeling in Plant Development.** *Annu Rev Cell Dev Biol.* **18:** 707-746 (2002).

Goto, K *et al.* **Turning floral organs into leaves, leaves into floral organs**. *Curr. Opin. Genet. Dev.* **11**:449-456 (2001).

Grandori, C *et al.* **The Myc/Max/Mad network and the transcriptional control of cell behaviour**. *Annu. Rev. Cell Dev. Biol.* **16:** 653-699. (2000).

Grimanelli, D *et al.* **Developmental genetics of gametophytic apomixis.** *Trends Genet.* **17(10):** 597-604. (2001).

Grossniklaus, U and Schneitz, K. **The Molecular and genetic basis of ovule and megagametophyte development**. *Cell & Dev. Biol.* **9:** 227-238 (1998a)

Grossniklaus, U *et al.* **Maternal Control of Embryogenesis by MEDEA, a Polycomb Group Gene in *Arabidopsis***. *Science.* **280**: 446-450 (1998b).

Grossniklaus, U *et al.* **Genomic imprinting and seed development: endosperm formation with and without sex.** *Curr. Opin.Plant Biol.* **4**: 21-27. (2001).

Gutjahr, T *et al.* **The polycomb-group gene, *extra sex combs*, encodes a nuclear member of the WD-40 repeat family**. *The EMBO Journal.* **14**:#17:4296-4306. (1995).

Haig, D. and Westoby, M. **Parent specific gene expression and the triploid endosperm.** *Am.Nat.* **134:** 147-155 (1989).

Haig, D. and Westoby, M. **Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis.** *Phil. Trans. R. Soc. Lond.* B **333**: 1-13 (1991).

Haseloff J, *et al.* **Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly.** *Proc. Natl. Acad. Sci. USA.* **94 (6):** 2122-2127 (1997).

Heim, MA *et al.* **The Basic Helix-Loop-Helix Transcription Factor Family in Plants: A Genome-Wide Study of Protein Structure and Functional Diversity.** *Mol. Biol. Evol.* **20(5):**735-747 (2003).

van Hengel, AJ *et al.* **Expression Pattern of the Carrot EP3 Endochitinase Genes in Suspension Cultures and in Developing Seeds.** *Plant Phys.* **117 (1):** 43 53 (1998).

Hennig, L *et al.* ***Arabidopsis*** **MSI1 is required for epigenetic maintenance of reproductive development.** *Development.* **130:** 2555-2565 (2003).

Holdeman, R *et al.* **MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a *Drosophila* Polycomb group protein.** *Development.* **125:**2457-2467 (1998).

Hong, SK *et al.* **How is embryo size genetically regulated in rice?** *Development.* **122**: 2051-2058 (1996).

Honma, T and Goto, K. **Complexes of MADS-box proteins are sufficient to convert leaves into floral organs.** *Nature.* **409:** 525 - 529 (2001).

Jenik, PD and Irish, VF. **The *Arabidopsis* floral homeotic gene APETALA3 differentially regulates intercellular signaling required for petal and stamen development.** *Development.* **128:** 13-23 (2001).

Jones CA, *et al.* **The Drosophila ESC and E(Z) proteins are direct partners in Polycomb group-mediated repression.** *Mol Cell Biol.* **18:**2825-2834 (1998).

Kaufman, PD *et al.* **Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I.** *Genes Dev.* **11:** 345-357 (1997).

Kehle, J *et al.* **dMi-2, a hunchback-interacting protein that functions on Polycomb repression.** *Science.* **282**:1897-1900 (1998).

Kelly, WG *et al.* **Chromatin silencing and the maintenance of a functional germline in *Caenorhabditis elegans*.** *Development.* **125:** 2451-2456 (1998).

Kenzior, AL and Folk, WR. **AtMSI4 and RbAp48 WD-40 repeat proteins bind metal ions.** *FEBS Lett.* **440:** 425-429 (1998).

Kindiger, B *et al.* **Assignment of a gene(s) conferring apomixis in *Tripsacum* to a chromosome arm: Cytological and molecular evidence.** *Genome.* **39 (6):** 1133-1141 (1996).

Kinoshita, T *et al.* **Imprinting of the *MEDEA* Polycomb Gene in the *Arabidopsis* Endosperm.** *The Plant Cell.* **11**: 1945-1952 (1999).

Kinoshita, T *et al.* **Polycomb repression of flowering during early plant development.** *PNAS.* **98 (24):** 14156-14161 (2001).

Kiyosue, T *et al.* **Control of fertilization-independent endosperm development by the *MEDEA* polycomb gene in *Arabidopsis*.** *Proc. Natl. Acad. Sci. USA.* **96**: 4186-4191 (1999).

Klekowski, EJ *et al.* **Mangrove genetics 4. Postzygotic mutations fixed as periclinal chimeras.** *Int. J. Plant Sci.* **157**:398-435 (1996).

Kohalmi, SE, *et al.* **A practical guide to using the yeast 2-hybrid system**. In *Differentially Expressed Genes in Plants*, E. Hansen and G. Harper, eds (London: Taylor and Francis), pp. 63–82. (1997).

Kohalmi, SE *et al.* **Identification and Characterisation of protein interactions using the yeast 2-hybrid system.** *Plant Molecular Biology Manual MI* 1-30 (1998).

Köhler, C and Grossniklaus, U. **Epigenetic inheritance of expression states in plant development: the role of *Polycomb* group proteins**. *Curr. Opin. Cell Biol.* **14**:773-779 (2002).

Köhler, C *et al.* **The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1**. *Genes Dev.* **17**:1540-1553 (2003).

Köhler, C *et al.* **Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development.** *EMBO J.* **22 (18):** 4804-4814 (2003).

Koltunow, AM *et al.* **Apomixis: Molecular Strategies For The Generation Of Genetically Identical Seeds Without Fertilization.** *Plant Phys.* **108 (4):** 1345-1352 (1995).

Koorneef, M *et al.* **The isolation and characterization of abscisic acid-insensitive mutants of *Arabidopsis thaliana*.** *Physiol Plant.* **61:** 377-383 (1984).

Korf I, *et al.* **The Polycomb group in Caenorhabditis elegans and maternal control of germline development**. *Development*. **125:**2469-2478 (1998).

Kowalski, SP *et al.* **Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization**. Genetics. **138:** 499-510 (1994).

Kranz, E and Dresselhaus, T. ***In vitro* fertilization with isolated higher plant gametes.** *Trends in Plant Sci*. **1 (3):** 82-89 MAR (1996).

Kuzmichev, A *et al.* **Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein.** *Genes Dev*. 16:2893-2905 (2002).

LaJeunesse, D and Shearn, A. **E(z) a *Polycomb* group gene or a *trithorax* group gene?** *Development*. **122:** 2189-97 (1996).

Lawrence, PA *et al.* **Homeobox genes: their function in *Drosophila* segmentation and pattern formation**. *Cell* **78**:181-189 (1994).

Lessard, J *et al.* **Functional antagonism of the polycomb-group genes eed and Bmi1 in hemopoietic cell proliferation**. *Genes Dev.* **13**:2691-2703 (1999).

Lin X *et al.* **Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana.*** *Nature*. **402:**761-768 (1999).

Long, D *et al.* **The maize transposable element system *Ac/Ds* as a mutagen in *Arabidopsis*: Identification of an *albino* mutation induced by *Ds* insertion.** *Proc.Natl. Acad.Sci. U.S.A.* **90:** 10370-10374 (1993).

Lotan T, *et al.* ***Arabidopsis* LEAFY COTYLEDON 1 is sufficient to induce embryo development in vegetative tissue.** *Cell.* **93**: 1195-1205 (1998).

Lu, XW and Horvitz, HR. **lin-35 and lin-53, two genes that antagonize a *C. elegans* Ras pathway, encode proteins similar to Rb and its binding protein RbAp48.** *Cell.* **95:** 981-991 (1998).

Luo, M *et al.* **Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*.** *PNAS. USA.* **96**:296-301 (1999).

Luo, M *et al.* **Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing *Arabidopsis* seeds.** *PNAS.* **97**: (19) 10637-10642 (2000).

Lyko, F and Paro, R. **Chromosomal elements conferring epigenetic inheritance.** *Bioessays.* **21**:824-832 (1999).

Mable, BK and Otto, SP. **The evolution of life cycles with haploid and diploid phases.** *Bioessays.* **20:**453-462 (1998).

Mak, W *et al.* **Mitotically stable association of polycomb group proteins eed and enx1 with the inactive x chromosome in trophoblast stem cells.** *Curr. Biol.* **1**:1016-1020 (2002).

Massari, ME, and Murre, C. **Helix-loop-helix proteins: Regulators of transcription in eukaryotic organisms.** *Mol. Cell. Biol.* **20:** 429-440. (2000).

Mayer, K *et al.* **Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.** *Nature*. **402:** 769-777 (1999).

McGrath, JM *et al.* **Duplicate sequences with similarity to expressed genes in the genome of Arabidopsis thaliana**. *Theor. Appl. Genet.* **86:**880-888 (1993).

Mizukami, Y and Ma, H. **Ectopic expression of the floral homeotic gene AGAMOUS in transgenic *Arabidopsis* plants alters floral organ identity.** *Cell.* 71:119-131 (1992).

Mizukami, Y and Ma, H. **Determination of Arabidopsis floral meristem identity by AGAMOUS.** *Plant Cell.* 9:393-408.

Moffat, AS. **For Plants, Reproduction Without Sex May Be Better**. *Science*. **294**:2463-2465 (2001).

Moon, YH *et al.* **EMF Genes Maintain Vegetative Development by Repressing the Flower Program in Arabidopsis.** *Plant Cell.* **15**:681-693 (2003).

Mora-Garcia, S *et al.* **Genomic imprinting: Seeds of conflict.** *Curr. Biol.* **10**:R71-R74. (2000).

Müller, J *et al.* **Histone Methyltransferase Activity of a *Drosophila* Polycomb Group Repressor Complex**. *Cell.* **111**:197-208 (2002).

Nambara, E *et al.* **A mutant of *Arabidopsis* which is defective in seed development and storage protein accumulation is a new *abi3* allele.** *Plant J.* **2:** 435-441 (1992).

Neer, MJ *et al.* **The ancient regulatory-protein family of WD-repeat proteins.** *Nature.* **371**:297-300 (1994).

Ng, J *et al.* **Evolutionary conservation and Predicted Structure of the *Drosophila* extra sex combs Repressor Protein.** *Mol. Cell. Biol.* **17:** 6663-6672. (1997).

Ng, J *et al.* **A *Drosophila* ESC-E(z) Protein Complex Is Distinct from Other Polycomb Group Complexes and Contains Covalently Modified ESC.** *Mol. Cell.Biol.* **20**: 3069-3078 (2000).

Nicolas E, *et al.* **RbAp48 belongs to the histone deacetylase complex that associates with the retinoblastoma protein.** *J. Biol. Chem.* **275(13):**9797-804 (2000).

Normark, BB. **The evolution of alternative genetic systems in insects.** *Annu. Rev. Entomol.* **48:**397-423 (2003).

Notredame, C *et al.* **T-Coffee: A novel method for multiple sequence alignments.** *J. Mol. Biol.* **302**: 205-217 (2000).

O'Carroll *et al.* **The polycomb-group gene Ezh2 is required for early mouse development.** *Mol. Cell. Biol.* **21**:4330-4336 (2001).

Ohad, N *et al.* **A mutation that allows endosperm development without fertilization.** *PNAS. USA.* **93**:5319-5324 (1996).

Ohad, N *et al.* **Mutations in *FIE*, a WD Polycomb Group gene, Allow Endosperm Development without Fertilization.** *Plant Cell.* **11**: 407-415 (1999).

Orlando, V *et al.* **Binding of Trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early *Drosophila* embryogenesis.** EMBO J. **17:** 5141-51150. (1998).

Orlando, V. **Polycomb, Epigenomes, and Control of Cell Identity**. *Cell.* **112**: 599-606 (2003).

Otte, AP and Kwaks, THJ. **Gene repression by Polycomb group protein complexes: a distinct complex for every occasion?** *Curr. Opin.Gen.Devt.* **13(5):** 448-454 (2003).

Pal-Bhadra, M *et al.* **RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in *Drosophila*.** *Mol. Cell.* **9:** 315-327 (2002).

Paro, R. **Propagating memories of transcriptional states**. *Trends in Genetics.* **11**: 295-7. (1995).

de Pater, S *et al.* **RAP-1 is an *Arabidopsis* MYC-like R protein homologue, that binds to G-box sequence motifs**. *Plant Mol. Biol.* **34**: 169-174 (1997).

Paterson AA, *et al.* **Towards a unified genetic map of higher plants transcending the monocot dicot divergence.** *Nat. Genet.* **14:**380-382 (1996).

Pirrotta, V. **Polycombing the Genome: PcG, trxG, and chromatin silencing.** *Cell.* **93:** 333-336. (1998).

Pelaz, S *et al.* **B and C floral organ identity functions require** SEPALLATA **MADS-box genes.** *Nature*. **405 (6783):** 200-203 (2000).

Pollok, BA *et al.* **Using GFP in FRET-based applications.** *Trends in Cell Biol.* **9**:57-60. (1999).

Ptashne, M. **A Genetic Switch: Phage Lambda and Higher Organisms**. Blackwell Science, Inc. (1996).

Qian YW and Lee EY. **Dual retinoblastoma-binding proteins with properties related to a negative regulator of ras in yeast.** *J Biol Chem.***270(43):**25507-13 (1995).

Que, Q *et al.* **The frequency and degree of cosuppression by sense chalcone synthase transgenes are dependent on transgene promoter strength and are reduced by premature nonsense codons in the transgene coding sequence.** *Plant Cell.* **9**:1357–1368 (1997).

Ray, S *et al.* **Maternal effects of the short integument mutation on embryo development in** *Arabidopsis. Dev. Biol.* **180 (1):** 365-369 (1996).

Rea, S *et al.* **Regulation of chromatin structure by site-specific histone H3 methyltransferases.** *Nature*. **406**:593-599 (2000).

Riechmann, JL and Ratcliffe, OJ. **A genomic perspective on plant transcription factors.** *Curr. Opin. Plant Biol.* **3**:423-434 (2000).

Ringrose, L *et al.* **Genome-Wide Prediction of Polycomb/Trithorax Response Elements in Drosophila melanogaster.** *Dev. Cell.* **5**:759-771 (2003).

Rotman, N *et al.* **Female Control of Male Gamete Delivery during Fertilization in *Arabidopsis thaliana.*** *Curr. Biol.* **13**:*5*:432-436 (2003).

Sadowski, I *et al.* **GAL4-VP16 is an unusually potent transcriptional activator.** *Nature.* **335**: 563 - 564 (1988).

Schneitz, K. **The molecular and genetic control of ovule development**. *Curr. Opin Plant Biol.* **2**:13-17 (1999).

Schumacher, A *et al.* **Positional cloning of a global regulator of anterior-posterior patterning in mice**. *Nature*. **348**:648 (1996).

Schumacher, A *et al.* **The Murine polycomb-Group Gene eed and Its Human Orthologue: Functional Implications of Evolutionary conservation**. *Genomics.* **54**: 79-88. (1998).

Scott, RJ *et al.* **Medea: murder or mistrial?** *Trends In Plant Sci.* **3**: 460-461. (1998).

Scott, RJ *et al.* **Parent-of-origin effects on seed development in Arabidopsis thaliana.** *Development.* **125:** 3329-3341 (1998).

Sewalt, RGAB *et al.* **Characterization of Interactions between the Mammalian Polycomb-Group Proteins Enx1/EZH2 and EED Suggests the Existence of Different Mammalian Polycomb-Group Protein Complexes.** *Mol. Cell. Biol.* **18**: 3586-3595 (1998).

Seydoux, G *et al.* **Launching the germline in *Caenorhabditis elegans*: regulation of gene expression in early germ cells.** *Development.* **126**: 3275-3283 (1999).

Shao, Z *et al.* **Stabilization of chromatin Structure by PRC1, a Polycomb complex.** *Cell.* 98: 37-46. (1999).

Sijen, T *et al.* **Transcriptional and posttranscriptional gene silencing are mechanistically related.** *Curr. Biol.* **11**:436–440 (2001).

Simon *et al.* **The extra sex combs product contains WD40 repeats and its time of action implies a role distinct from other Polycomb group products.** *Mech. Dev.* **53**: 197-208. (1995).

Smolen, GA *et al.* **Dominant Alleles of the Basic Helix-Loop-Helix Transcription Factor ATR2 Activate Stress-Responsive Genes in Arabidopsis.** *Genetics.* **161**: 1235-1246 (2002).

Sondek, J *et al.* **Crystal structure of a G_A protein βγ dimer at 2.1Å resolution.** *Nature* **379**:369-374. (1996)

Sørensen, MB *et al.* **Polycomb group genes control pattern formation in plant seed.** *Curr. Biol.* **11**: 277-281 (2001).

Sørensen, MB *et al.* **Cellularisation in the endosperm of *Arabidopsis thaliana* is coupled to mitosis and shares multiple components with cytokinesis.** *Development.* (2002).

Spillane, C *et al.* **Interaction of the Arabidopsis Polycomb group proteins FIE and MEA mediates their common phenotypes.** *Curr. Biol.* **10**:1535-1538 (2000).

Springer, NM *et al.* **Sequence Relationships, Conserved Domains, and Expression Patterns for Maize Homologs of the Polycomb Group Genes *E(z), esc* and *E(Pc).*** *Plant Phys.* **128:** 1332-1345 (2002).

Springer, P and Holding, DR. **The Arabidopsis gene PROLIFERA is required for proper cytokinesis during seed development.** *Planta. 214*: 373-382. (2002).

Steeves, TA and Sussex, IM. **Patterns in plant development.** Chpt. 17. Cambridge, UK: Cambridge Univ. Press (2002).

Stam, M *et al.* **Position-dependent methylation and transcriptional silencing of transgenes in inverted T-DNA repeats: implications for posttranscriptional silencing of homologous host genes in plants.** *Mol. Cell. Biol.* **18:** 6165–6177 (1998).

Stotz, A and Linder, P. **The *ADE2* gene from *Saccharomyces cerevisiae*: sequence and new vectors.** *Gene.* **95**: 91–98 (1990).

Sussman, MR *et al.* **The *Arabidopsis* knockout facility at the University of Wisconson-Madison.** *Plant Physiology.* **124**:1465-1467 (2000).

Tanaka H, *et al.* **A subtilisin-like serine protease is required for epidermal surface formation in Arabidopsis embryos and juvenile plants.** *Development.* **128 (23):** 4681-4689 (2001).

Taylor, RL. **The foliar embryos of *Malaxix paludosa.*** *Can J Bot.* **45:** 1553-1556 (1967).

Terryn, N *et al.* **Evidence for an ancient chromosomal duplication in *Arabidopsis thaliana* by sequencing and analysing a 400-kb contig at the APETALA2 locus on chromosome 4.** *FEBS Lett.* **445:**237-245 (1999).

Thompson, JD *et al.* **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice.** *Nuc. Acids Res.* **22**:4673-4680 (1994).

Tie, F *et al.* **The *Drosophila* Polycomb Group proteins ESC and E (z) bind directly to each other and co-localize at multiple chromosomal sites.** *Development.* **125**:3483-3496 (1998).

Tie, F *et al.* **The *Drosophila* Polycomb group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3.** *Development.* **128**:275-286 (2001).

Toledo-Ortiz G, *et al.* **The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family.** *Plant Cell.* **15**:1749-1770 (2003).

Tyler JK, *et al.* **The p55 subunit of Drosophila chromatin assembly factor 1 is homologous to a histone deacetylase-associated protein.** *Mol Cell Biol.* **16(11):**6149-59 (1996).

Tzung-Fu, H *et al.* **From flour to flower: how Polycomb group proteins influence multiple aspects of plant development.** *Trends in Plant Sci.* **8(9):**439-45 (2003).

van Lohuizen , M *et al.* **Interaction of Mouse Polycomb-Group (Pc-G) Proteins Enx1 and Enx2 with Eed: Indication for separate Pc-G complexes.** *Mol. Cell. Biol.* **18**: 3572-3579 (1998).

van der Vlag, J *et al.* **Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation.** *Nature Genetics.* **23**: 474-477 (1999).

Vaucheret, H *et al.* **A transcriptionally active state is required for post-transcriptional silencing (cosuppression) of nitrate reductase host genes and transgenes.** *Plant Cell* **9:**1495–1504 (1997).

Vielle-Calzada, J-P *et al.* **Maintenance of Genomic Imprinting at the *Arabidopsis* media Locus Requires Zygotic DDM1 Activity.** *Gen.. & Dev.* **13 (22):** 2971-2982 (1999).

Vielle-Calzada, J-P *et al.* **Delayed activation of the paternal genome during seed development.** *Nature.* **404:** 91-94. (2000).

Vinkenoog, R *et al.* **Hypomethylation Promotes Autonomous Endosperm Development and Rescues Postfertilization Lethality in *fie* Mutants.** *Plant Cell.* 12:2271-2282 (2000).

Waizenegger, I *et al.* **The *Arabidopsis KNOLLE* and *KEULE* genes interact to promote vesicle fusion during cytokinesis.** *Curr Biol.* **10(21):**1371-4 (2000).

Wall, MA *et al.* **The structure of the G protein heterotrimer $G_{i\alpha 1}\beta_1\gamma_2$.** *Cell.* **83**:1047-1058 (1995).

Wang, J *et al.* **Imprinted X inactivation maintained by a mouse polycomb group gene.** *Nat. Genet.* **28**:371-375 (2001).

Waterhouse, PM *et al.* **Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA.** *PNAS. USA.* **95:**13959–13964 (1998).

Weigel, D and Meyerowitz, EM. **The ABCs of Floral Homeotic Genes.** *Cell.* **78:** 203-209 (1994).

Weijers, D *et al.* **Early paternal gene activity in *Arabidopsis*.** *Nature*. **414**:709-710 (2001).

Williams, EG and Maheswaran, G. **Somatic embryogenesis: factors influencing coordinated behaviour of cells as an embryogenic group.** *Annals of Botany.* **57 (4):** 443-462 (1986).

Xiao, W *et al.* **Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase.** *Dev Cell.* **5(6):** 891-901(2003).

Yadegari, R *et al.* **Mutations in the FIE and MEA Genes That Encode Interacting Polycomb Proteins Cause Parent-of-Origin Effects on Seed Development by Distinct Mechanisms**. *Plant Cell.* **12**: 2367-2381 (2000).

Yamaguchi-Shinozaki and Shinozaki. **The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of rd22, a gene responsive to dehydration stress in *Arabidopsis thaliana*.** *Mol. Gen. Genet.* **238**:17-25 (1993).

Yoshida, N *et al.* **EMBRYONIC FLOWER2 a Novel Polycomb Group Protein Homolog, Mediates shoot Development and Flowering in Arabidopsis.** *Plant Cell.* **13**:2471-2481 (2001).

Zhang, C-C *et al.* **Segmental determination in Drosophila conferred by hunchback (hb), a repressor of the homeotic gene Ultrabithorax (Ubx).** *PNAS. USA.* **89**:7511-7515. (1992).

Interaction of the *Arabidopsis* Polycomb group proteins FIE and MEA mediates their common phenotypes

C. Spillane*†‡, C. MacDougall‡§, C. Stock‡§, C. Köhler*†, J-P. Vielle-Calzada#¶, S.M. Nunes†, U. Grossniklaus*†# and J. Goodrich§

Genes of the *FERTILISATION INDEPENDENT SEED* (*FIS*) class regulate cell proliferation during reproductive development in *Arabidopsis* [1–5]. The *FIS* genes *FERTILISATION INDEPENDENT ENDOSPERM* (*FIE*) and *MEDEA* (*MEA*) encode homologs of animal Polycomb group (Pc-G) proteins, transcriptional regulators that modify chromatin structure and are thought to form multimeric complexes [3–11]. To test whether similarities in *fis* mutant phenotypes reflect interactions between their protein products, we characterised *FIE* RNA and protein localisation *in vivo*, and FIE protein interactions in yeast and *in vitro*. Expression of *FIE* mRNA overlaps with that of *MEA* during embryo sac and seed development and is unaffected in *mea* mutants. Results from the yeast two-hybrid system and an *in vitro* pull-down assay indicate that MEA and FIE proteins interact. The relevance of this interaction *in vivo* is supported by the finding that FIE and MEA co-localise in the nucleus in transfected plant cells. Interaction of MEA and FIE is mediated by the amino-terminal region of MEA. Despite sequence divergence in this domain, MEA can interact with its corresponding animal partner Extra sex combs (ESC) in the yeast two-hybrid system. We conclude that FIE and MEA act together as part of a multimeric complex and that this accounts for the similarities in mutant phenotypes. We propose that an ancient mechanism for chromatin modification has been independently recruited to different developmental processes in the two kingdoms.

Addresses: *Institute of Plant Biology, University of Zürich, CH-8008 Zürich, Switzerland. †Friedrich Miescher Institute, CH-4058 Basel, Switzerland. §Institute of Cell and Molecular Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JH, UK. #Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

Present address: ¶CINVESTAV-Irapuato, Plant Biotechnology Unit, CP 36 500, Irapuato GTO, Mexico.

‡These authors contributed equally to this work.

Correspondence: Justin Goodrich
E-mail: Justin.Goodrich@ed.ac.uk

Received: **25 August 2000**
Revised: **5 October 2000**
Accepted: **5 October 2000**

Published: **17 November 2000**

Current Biology 2000, 10:1535–1538

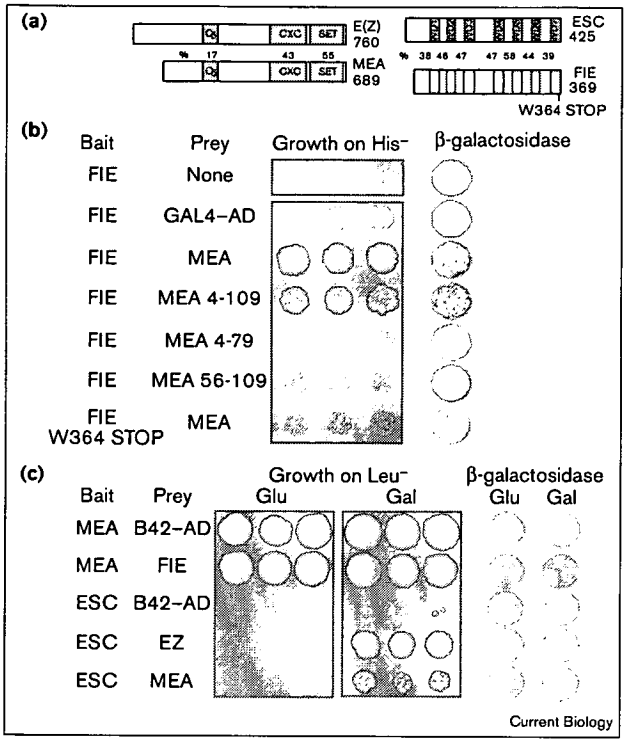
0960-9822/00/$ – see front matter

Results and discussion

MEA and *FIE* encode homologues of the *Drosophila* Enhancer of zeste E(Z) and ESC proteins, respectively (Figure 1a; see also [3–7] and Supplementary material). These are widely conserved and have been shown to co-localize in complexes distinct from the other Pc-G proteins [9–16]. To test whether plant MEA and FIE could interact, yeast two-hybrid assays were performed. The near full-length proteins were expressed as 'bait' fusions to the yeast Gal4 DNA-binding (DB) domain, and as 'prey' fusions to the Gal4 transcriptional activation (TA) domain [17] (see Supplementary material). Yeast strains that carried the FIE bait construct alone were unable to grow in the absence of histidine and did not express β-galactosidase, indicating that the DB–FIE fusion did not activate either reporter gene. In addition, GAL4–TA and FIE–DB did not interact in yeast. However, yeast transformants carrying *DB–FIE* and *TA–MEA* expressed both reporters, suggesting that FIE and MEA interact (Figure 1b). We did not detect a reciprocal interaction when FIE was expressed as a prey fusion and MEA as a bait (data not shown), presumably because one or both fusions were non-functional. In an alternative two-hybrid system [18], MEA–DB constructs on their own were able to weakly activate reporter gene expression, but gave increased β-galactosidase activity in the presence of FIE–TA prey (Figure 1c), also consistent with an FIE–MEA interaction. To test whether mutations that eliminate FIE activity *in planta* also affect its interaction with MEA, we introduced the *fie-6* mutation into the *FIE–GAL4–DB* bait. *fie-6* confers a strong mutant phenotype equivalent to that of alleles encoding extensive truncations or deletions [7]. It encodes FIE W364STOP, which truncates the FIE protein by six residues and disrupts the highly conserved WD residues at the carboxy-terminal end of WD repeat motif seven, which are predicted to stabilise the β-propeller structure (for example [8,9,19]). Introduction of the W364STOP mutation into the FIE bait prevented its interaction with MEA (Figure 1b), consistent with the FIE–MEA interaction being necessary for FIE activity *in planta*.

To map the region of the MEA protein required for interaction with FIE, a series of carboxy-terminal truncations of MEA were expressed as fusions to the Gal4 activation domain. We found that the amino-terminal residues 4–109 were sufficient for interaction with FIE. Proteins with more extensive deletions (residues 4–79) were unable to interact with FIE (Figure 1b) and a series of amino-terminal

Figure 1



FIE and MEA interact in yeast. **(a)** Alignment of MEA and E(Z), and FIE and ESC proteins. Amino acid identities in conserved domains are indicated. C_5, CYS and SET, conserved cysteine-rich regions and SET domain previously described (for example [4]); WD, WD repeat motif. **(b)** Assays using a Gal4-based yeast two-hybrid system [16]. Three independent transformants are shown for each growth assay on His$^-$ media. For β-galactosidase activity assays, six independent transformants were grown on filter paper, lysed, and incubated on media containing X-Gal (see Supplementary material). **(c)** Assays using a LexA–DB-based system [18]. In this system, *LEU2* and *lacZ* (β-galactosidase) reporter genes are transcriptionally dependent on LexA. Expression of the prey construct is galactose inducible, so colonies were grown on glucose- or galactose-containing media to compare reporter gene activation in the presence of bait alone or bait with prey, respectively. Three independent transformants are shown for each assay. E(Z)–DB and MEA–DB alone activated the *LEU2* reporter but gave weaker activation of the β-galactosidase reporter, FIE–DB alone strongly activated both reporters and therefore could not be used in two-hybrid assays.

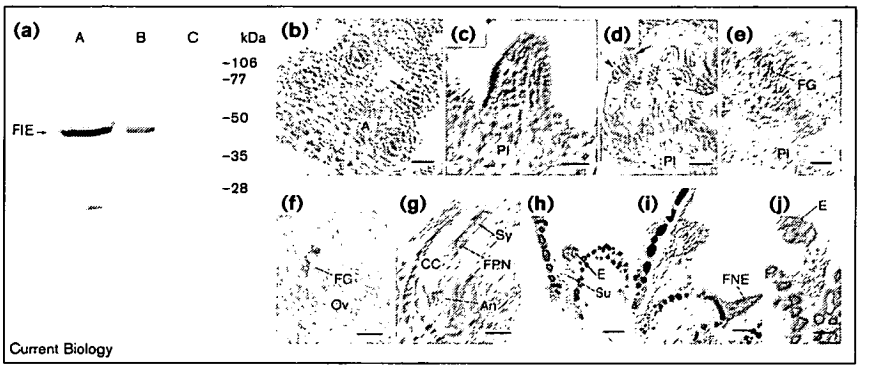
deletions all tested negative for interaction with FIE. Together, these results suggested that residues 4–109 of MEA were sufficient for interaction with FIE.

An amino-terminal portion of the *Drosophila* E(Z) protein is sufficient for its interaction with ESC [8–10]. This suggested that the interaction between E(Z) and ESC homologues is conserved between plants and animals, and is mediated by amino-terminal regions of E(Z) and MEA. However, there is little sequence similarity between these regions of MEA and E(Z) indicating that the domains mediating the interaction have diverged in the plant and animal lineages. To test whether the animal and plant counterparts could recognise one another in two-hybrid assays (Figure 1c), we used a LexA-DB-based system in which E(Z) and ESC interaction had been demonstrated [9]. In this system, Leu2 and β-galactosidase reporters are transcriptionally dependent upon LexA [18]. TA–E(Z) and TA–MEA preys both interacted with DB–ESC bait and activated the reporters, although the interaction with MEA was weaker than with E(Z) (Figure 1c). Thus MEA and EZ can both interact with ESC, despite their divergent sequences in the amino-terminal region.

To confirm independently that FIE and MEA interact, we performed *in vitro* binding assays (Figure 2a). Full-length MEA protein was expressed as a glutathione-*S*-transferase (GST) fusion in *Escherichia coli*, purified, and immobilized on glutathione–agarose beads. To test for binding to FIE, extracts were prepared from *E. coli* cells that expressed FIE tagged with Xpress epitope and incubated with equal volumes of beads attached to GST or GST–MEA. Western blots of the bead-associated proteins were probed with anti-Xpress antibodies. As shown in Figure 2, Xpress–FIE protein bound to GST–MEA (lane B) but not to GST alone (lane C), suggesting a direct physical interaction between MEA and FIE proteins. Thus, the MEA and FIE proteins bind to each other *in vitro* as well as in yeast.

If the physical interaction of *FIE* and *MEA* is biologically relevant, their products would be expected to be coexpressed *in vivo*. The *in vivo* expression pattern of *MEA* was described previously [20] and is consistent with its roles in the embryo sac and during zygotic development: before fertilisation *MEA* RNA is expressed in the synergids, the egg and central cell of the embryo sac; after fertilisation, *MEA* RNA is present in both fertilization products of the developing seed, the embryo and endosperm. To determine localisation of *FIE* mRNA *in vivo*, we hybridised digoxigenin-labelled *FIE* antisense (Figure 2b-j) and sense RNA probes (see Supplementary material) to tissue sections of developing reproductive organs *in situ*. *FIE* is initially expressed in the developing anther and the young ovule primordium. In the anther, *FIE* mRNA is abundant in microspore mother cells undergoing meiosis, in microsporocytes and in the tapetum (Figure 2b), but absent from vascular bundles, the connective tissue and the filament. *FIE* mRNA is also absent from pollen grains at subsequent developmental stages (data not shown). In the female reproductive organs, *FIE* mRNA is initially present in all cells of the young ovule primordium before archesporial differentiation (Figure 2c). *FIE* is not expressed in the placenta or the developing carpel. It is abundantly expressed in the ovule sporophytic tissue and the megaspore mother cell before meiosis (Figure 2d). Subsequently, at the initiation of megagametogenesis, *FIE* mRNA levels appear to decrease in the nucellus and integuments but to increase in the developing embryo sac (megagametophyte).

Figure 2



(a) *In vitro* binding of MEA and FIE. Bacterial extract containing Xpress–FIE protein was tested for binding to GST–MEA (lane B) or GST (lane C). Proteins that bound to GST or GST–MEA were separated by SDS–PAGE, transferred to PVDF membrane, and incubated with anti-Xpress antibodies. The input lane (lane A) contains 1.5% of the volume of bacterial extract used in the binding assay. **(b–j)** Localisation of *FIE* mRNA (shown by brown staining) in reproductive organs and developing seeds. (b) Young stamens; *FIE* mRNA is localised in the lodicule (arrow) but not the rest of the anther. (c) Young ovules before megaspore mother cell differentiation. (d) Ovules undergoing meiosis; the arrow indicates the megaspore mother cell and the arrowhead the integument primordia. (e) Two-nucleate female gametophyte. (f) Eight-nucleate female gametophyte. (g) Mature female gametophyte. (h) Seed containing an embryo at the octant stage. (i) Seed with an embryo at the transition stage. (j) Developing *mea/mea* seeds. The dark brown cells in (h,i,j), are endothelial cells, which commonly exhibit strong background staining. Control hybridisations using sense FIE probes are presented in Supplementary material. A, anther; An, antipodals; CC, central cell; E, embryo proper; FG, female gametophyte; FNE, free nuclear endosperm; FPN, fused polar nuclei; PT, placental tissue; Su, suspensor. Scale bars represent (b) 12 μm; (c) 8 μm; (d) 22 μm; (e) 15 μm; (f,g) 10 μm; (h,i) 30 μm; (j) 25 μm.

At the two-nucleated stage, *FIE* mRNA is abundant in the developing megagametophyte (Figure 2e). At maturity, expression is restricted to all cells of the megagametophyte, the endothelium and the persistent nucellar cells located at the chalazal region of the ovule (Figure 2f,g). After fertilisation, during early seed development, *FIE* is expressed in both embryo and endosperm, but not in the developing seed coat. At the octant stage of embryo development, *FIE* mRNA is localised in the embryo proper but not the suspensor (Figure 2h,i). In the developing endosperm, *FIE* mRNA is weakly expressed in free endosperm nuclei covering the mycropylar region of the embryo sac, but is abundant in large free nuclear nodules present in the chalazal region of the developing seed (Figure 2i). *FIE* mRNA is absent from the seed coat or other maternal tissues from developing siliques.

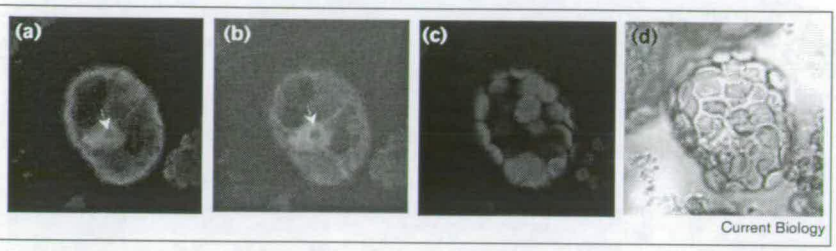
The distribution of *FIE* mRNA in the megagametophyte and in developing seeds is therefore consistent with its biological function of repressing central cell proliferation before fertilisation, and controlling early embryo/endosperm development. Furthermore, the expression patterns of *FIE* and *MEA* [20] overlap in the central, egg and synergid cells of the embryo sac, and during early embryo and endosperm development (except for the suspensor), indicating the potential for their protein products to interact.

To test whether *MEA* might regulate *FIE* transcription, we characterised *FIE* expression in developing siliques of self-pollinated homozygous *mea/mea* plants (produced by embryo rescue from *mea* seed as described in [20]). We compared wild-type and *mea* mature ovules before and after fertilisation. For all stages of development examined, no differences were observed for the patterns of *FIE* mRNA accumulation (Figure 2j). This suggests that *MEA* does not regulate *FIE* at the transcriptional level.

To investigate the intracellular localisation of MEA and FIE proteins we performed transient expression experiments in cowpea protoplasts, using reporter gene constructs that expressed MEA or FIE as amino-terminal fusions to yellow fluorescent protein (YFP) or cyan fluorescent protein (CFP). Cotransfections of cowpea protoplast cells with *FIE–CFP* (Figure 3a) and *MEA–YFP* (Figure 3b) indicated that FIE and MEA co-localised to the nucleus and were both consistently excluded from the nucleolus. Nuclear localisation of both MEA and FIE was also observed in single transfections and for the reciprocal constructs FIE–YFP and MEA–CFP. The *in vivo* co-localisation of MEA and FIE proteins in the nucleus of plant cells, in addition to the demonstration of their physical interactions in yeast and *in vitro* (Figures 1,2), is consistent with a MEA–FIE Pc-G protein complex occurring *in vivo*.

The *FIE* and *MEA* genes confer similar mutant phenotypes in both unfertilized embryo sacs and developing seeds; that is, fertilisation-independent endosperm proliferation and the abortion of fertilised seeds that carry a maternally inherited mutant *mea* or *fie* allele. We have shown that MEA and FIE interact in yeast and *in vitro*, and that the *fie-6* mutation which drastically reduces *FIE* activity *in vivo* also eliminates interaction of FIE and MEA. Furthermore, *FIE* and *MEA* RNA expression overlaps *in vivo* and the proteins co-localise within the nucleus *in planta*. We conclude that the similarity of the *fie* and *mea* phenotypes reflects the interaction of their gene products

Figure 3



Cowpea protoplast co-transfected with pFIE–CFP and pMEA–YFP shows co-localisation of FIE–CFP and MEA–YFP fusion proteins. **(a)** FIE–CFP; **(b)** MEA–YFP; **(c)** chloroplast autofluorescence; **(d)** phase-contrast microscopy. White arrows indicate the nucleolus. Co-localized FIE–CFP and MEA–YFP in the cytoplasm may be related to high levels of expression from the 35S promoter.

as part of a multiprotein complex. By analogy with animal systems [11] it is likely that the FIE–MEA complex modifies higher-order chromatin structure around their target genes to confer a transcriptionally inactive state.

The ESC and E(Z) proteins are more widely conserved than other members of the *Drosophila* and vertebrate Pc-G [12,13], suggesting that these two proteins are of more ancient origin than other Pc-G proteins. Because plant and animals are thought to have evolved multicellularity independently, an ancient Pc-G-dependent mechanism for chromatin modification has been recruited to independent developmental processes in the two kingdoms. Elucidation of the FIE/MEA regulatory network will test whether different target genes have been acquired in the two lineages.

Note added in proof

Luo *et al.* have also recently reported that FIE and MEA interact in two-hybrid assays [21].

Supplementary material

Supplementary material including a detailed description of Materials and methods and additional Results is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We thank Jeff Simon and Rick Jones for clones and information, Renate Schmidt for locating the *FIE* cDNA on the *Arabidopsis* physical map, Bill Crosby and Roger Brent for yeast two-hybrid vectors and technical advice and Santiago Mora-Garcia for constructs for FIE protein expression. We thank Khalid Shah, Dorus Gadella and Sacco de Vries (Wageningen Agricultural University) for invaluable assistance with the fluorescence microscopy analysis. U.G. acknowledges support by the Novartis Research Foundation, IAESTE, and the Kanton Zürich. C.K. and J-P. V-C. were supported by EMBO Long Term and Swiss National Science Foundation Fellowships, respectively, and U.G. is a Searle Scholar. J.G. is a Royal Society University research fellow. C.M.D. and C.S. were funded by a CAD award and a PhD studentship from the BBSRC, respectively.

References

1. Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ: **Fertilization-independent seed development in *Arabidopsis thaliana*.** *Proc Natl Acad Sci USA* 1997, **94**:4223-4228.
2. Ohad N, Margossian Y, Hsu Y-C, Williams C, Repetti P, Fischer RL: **A mutation that allows endosperm development without fertilisation.** *Proc Natl Acad Sci USA* 1996, **93**:5319-5324
3. Grossniklaus U, Vielle-Calzada JP: **...response: parental conflict and infanticide during embryogenesis.** *Trends Plant Sci* 1998, **3**:328
4. Grossniklaus U, VielleCalzada JP, Hoeppner MA, Gagliano WB: **Maternal control of embryogenesis by *MEDEA*, a Polycomb group gene in *Arabidopsis*.** *Science* 1998, **280**:446-450.
5. Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, *et al.*: **Control of fertilization-independent endosperm development by the *MEDEA* polycomb gene *Arabidopsis*.** *Proc Natl Acad Sci USA* 1999, **96**:4186-4191
6. Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM: **Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*.** *Proc Natl Acad Sci USA* 1999, **96**:296-301.
7. Ohad N, Yadegari R, Margossian L, Hannon M, Michaeli D, Harada JJ, *et al.*: **Mutations in *FIE*, a WD polycomb group gene, allow endosperm development without fertilization.** *Plant Cell* 1999, **11**:407-415
8. Tie F, Furuyama T, Harte PJ: **The Drosophila Polycomb Group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal sites.** *Development* 1998, **125**:3483-3496.
9. Jones CA, Ng J, Peterson AJ, Morgan K, Simon J, Jones RS: **The *Drosophila* ESC and E(Z) proteins are direct partners in Polycomb group-mediated repression.** *Mol Cell Biol* 1998, **18**:2825-2834.
10. Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu C-T, Bender W, Kingston RE: **Stabilisation of chromatin structure by PRC1, a Polycomb complex.** *Cell* 1999, **98**:37-46.
11. van der Vlag J, Otte AP: **Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation.** *Nat Genet* 1999, **23**:474-478.
12. Korf I, Fan Y, Strome S: **The Polycomb group in *Caenorhabditis elegans* and maternal control of germline development.** *Development* 1998, **125**:2469-2478.
13. Holdeman R, Nehrt S, Strome S: **MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a *Drosophila* Polycomb group protein.** *Development* 1998, **125**:2457-2467.
14. Van Lohuizen M, Tijms M, Voncken JW, Schumacher A, Magnuson T, Wientjens E: **Interaction of the mouse Polycomb-group (Pc-G) proteins Enx1 and Enx2 with Eed: indication for separate Pc-G complexes.** *Mol Cell Biol* 1998, **18**:3527-3579.
15. Ng J, Hart CM, Morgan K, Simon JA: **A *Drosophila* ESC-E(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC.** *Mol Cell Biol* 2000, **20**:3069-3078.
16. Sewalt R, van der Vlag J, Gunster MJ, Hamer KM, den Blaauwen JL, Satijn DPE, *et al.*: **Characterization of interactions between the mammalian polycomb-group proteins Enx1/E(Z)H2 and EED suggests the existence of different mammalian polycomb-group protein complexes.** *Mol Cell Biol* 1998, **18**:3586-3595.
17. Kohalmi SE, Reader LJV, Samach A, Nowak J, Haughn GW, Crosby WL: **Identification and characterization of protein interactions using the yeast two-hybrid system.** *Plant Mol Biol* 1998, **M1**:1-30.
18. Golemis EA, Gyuris J, Brent R: **Interaction trap/two-hybrid system to identify interacting proteins.** In *Current Protocols in Molecular Biology*. Edited by Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. John Wiley: New York; 1996:20.1.1–20-1-23.
19. Ng J, Li R, Morgan K, Simon J: **Evolutionary conservation and predicted structure of the *Drosophila* extra sex combs repressor protein.** *Mol Cell Biol* 1997, **17**:6663-6672.
20. VielleCalzada JP, Thomas J, Spillane C, Coluccio A, Hoeppner MA, Grossniklaus U: **Maintenance of genomic imprinting at the *Arabidopsis MEDEA* locus requires zygotic *DDM1* activity.** *Genes Dev* 1999, **13**:2971-2982.
21. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudhury A: **Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing *Arabidopsis* seeds.** *Proc Natl Acad Sci USA* 2000 **97**:10637-10642.